FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTCAGATCTGCTCGGTAGA CCTGGTGCACCACCACCATGTTGGCTGCAAGGCTGGTGTCTCCGGACACTACCTTCTAGG GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA GAACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT GATCAGATGGGAAGATGGTTTGTTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCCTCTTCTCATCAGAGCTGCATGGTACAC AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT AGTTCTTTTCAGCATGTTCCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT $oldsymbol{A}$ AGTGACTCAGCTTCTGGCTACATCAAATATCTTGTTTAATGGGGCAGATATGC ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT TTGGTGAATGTGAAAACTAAAGTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA GCTGAACTTAACAAAACTGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA ATAACTTTAAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG AATACAAACAGTATACTCATG

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCGTCTCCGCCTTCTGCAT TGTCAGTACTTTATGTCTGGGCCCAGCTGAACAGAGACATGATTGTATCATTTTGGTTTGGA ACACGATTTAAGGCCTGCTATTTACCCTGGGTTATCCTTGGATTCAACTATATCATCGGAGG CTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTTTCCTAATGTTCA GATACCCAATGGACTTGGGAGGAAGAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG CTGCCCAGTAGGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGCTGCTGATCAGAATGGCGGAGGCGGGAGACACAACTGGGGCCAGGGCTTTCGACTTGGAGACC AGTGAAGGGCGGCCTCGGGCAGCCTCTCAAGCCACATTTCCTCCCAGTGCTGGGTGCTGACACTTAACAACTGCGTTCTGGCTAACACTGTTGGACCTGACCACACTGAATGTAGTCTTTCAAGTACAACACAAGTTTCCTCACAAGTTTCACGAT TCTCATTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAATTGCAAAACTGAC TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT TCGTAGGTGGGCTTTTCCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACTTCTCATTTTGAATAGAGTTTGTGTG TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA TTTTTGTAAACTAATCCTTTTTATTGGTAAAAATTGTAAATTAAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC $\texttt{CCTTTGGGGCGGG} \underline{\textbf{ATG}} \texttt{GCTGCGGAAGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG}$ CGGGGTTCCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA AAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTACCT GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAAGAAAAA AACAGTTATCAGAGGCTAAAACAGAAGAGCCCACAGTGCATTCCAGTGAAGCTGCAATAATG CACAAAAAGGCCTGAAGATTCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAA CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAG<u>TA</u> **CTTACACTG**

FIGURE 6

MAAEEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVVSDLEHEEMKILREVL RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEEKQTLLKRRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAAATCAAAAGAGGAATATGACCAGGAA

FIGURE 8

TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG AGTGGA**ATG**GAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT CTGCCAATGAAGAAACAAGTATGATTATCTTCCAACTACTGTGAATGTGTGCTCAGAACTG AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTCATGAAGTGGTCCATTCCTGCCT TTCTTTATTTCCTGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATG GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAGGATAGTGCTGAA GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTTGTCTATTGTGGCCT TGACTGCCGGGACTAAAACTTTACAGCACAACTTGGCAGGACGTGGATTTCATCACGATGCC TACAGCAAAGGAATGGACTTTTCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTCAGTC ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA GAACAGCAAACTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTCAGTAGCC CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA TTTATTTATAATGCCAGCAAGCCTCAAGTTCCGGAATACGCACCTAGGCAAGAAAGGATCCG AGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTC**TAA**CTGGTACCCACATAGTTTGCA GCTCTCTTGAACCTTATTTTCACATTTTCAGTGTTTTGTAATATTTATCTTTTCACTTTGATA AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT TTTCCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT CTTCCCTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT TTCTCAGACACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAAAAAATCTTGAT AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC CAGATTGTCAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAAGGATCACTT AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTCATG GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT CATTCTCAGAAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG GTGGAAATTTGTAATTAAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFSNFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSGDGEELERLTKPKSDESDEDTF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293, 298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

FIGURE 11

CGGACGCGTGGGCGGACGCGTGGGCGGCCGGCTTGGCTAGCGCGCGGCGCCC GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG AGCGGCTCCGCGGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA GGTGAAGAAGAACCGAAAAAGAAGAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG CTATTGGATGTGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC CTGGGATGCCATCACAGACCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC TGGGTCGCCTTATGCCCTGGATCATCTTCTCCACGCCCCTGGCCGTCATTGCCTACTTCCTC ATCTGGTTCGTGCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTCATCAGCA ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCACTTCACAC AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTG TGCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGACCCTATGAAGCCCAGCAGTCTG AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTT ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC TACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT GGAACCGAGCCCATCTTCTCTCTCTTCTATGTCTTCACCAAGTTTGCCTCTGGAGTGTC ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG CTGGGCCTGCTCTTCAAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTCAGAAACAGACTCCACAG ${\tt AGCTGGCTAGCATCCTC}$ AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG TGGCCTCCTCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG SCPTSHTARPIGTCFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV ILILGVREQREPYEAQQSEPIAYFRGLRLVMSHGPYIKLITGFLFTSLAFMLVEGNFVLFCT YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKKAL QALRDEASSSGCSETDSTELASIL

FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATG**T GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT TTCATATTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCCAATGCTAAATATTGCGG CAGTTTTATGCATTGCTACCATTTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA GAGAACGTTATCATCAAATTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGG ACTTTCTATTGTGGCAAACTTCCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG TGCTTACCTTTGGTATGGGCTCATTATATGTTTGTTCAGACCATCCTTTCCTACCAAATG CAGCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG AGTAAGTGCACTTAGCATGCTGACTTGCTCATCAGTTTTTGCACAGTGGCAATTTTTGGGACTG ATTTAGAACAGAAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTTGGTTTTTTCCTGACTTACATTCGTGA TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATT**TGA**TGAAAGGAT AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTCACAGAAGTTGCTTA TTCTTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMQPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

FIGURE 16

FIGURE 17

CCCACGCGTCCGCCGCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGG $\tt CCGGGGTGCGGAGCCGAC \underline{\textbf{ATG}} \tt CGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC$ CTTCGCCTTGTACTTGCTGTCGACGCGACTGCCCCGCGGGGGGGAGACTGGGCTCCACCGAGG AGGCTGGAGGCAGGTCGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCCTGCTCTTCTGCGGCGCCCTA CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT TGCTACCTGCTCCCAGTATTTTTGGCAAACAGTTGGTGGTGTCCTACTTTCCTGATAAAGT TGAGACTTTTCCCCATGACACCAAACTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT CCCATCGTGCAGTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTCATCTGTGT GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA ${\tt CACA} {\color{red}{\bf TGA}} {\tt TCTGGATTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA$ TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT ACAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYR KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS IFGKQLVVSYFPDKVALLQRKVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPILNIPIVQFF FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ LNETSTANHIHSRKDT

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31.

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGAGGCCCGAGGGGGCCCCGCATGAATCATTTTT CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA ${\tt TAGGAAAATAACTTGGGATTTTATATTGGAAGAC} {\tt ATG} {\tt GATCTTGCTGCCAACGAGATCAGCA}$ TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC CCCCCGCAGTATCCTCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCCTGCC AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCCTTATCCATGGAGGAGACCTCTGAA CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCG ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT ${\tt GGAACCGCTTTCTCAGAACTG} {\color{red}{\bf TAG}} {\tt GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA}$ AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

FIGURE 20

MDLAANEISIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPFPYPWRRPLNRSQMLRELFPV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRHCQSVAMP
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

FIGURE 21

FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG CCACTGGTGCGCACGCTGCTAGACCGTGCCTATGAGCCGCTGGGGGCTGCAGTGGGGACTGCC CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAACTG GACTTCCTGGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT AGGCGATGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACCCCAGCAGCTCTCTGGAGTCGTGGAGTCGTCTTTGGCAGCTCTGGAGTCGAGTCTTTTGGCAGCTCTGGAGTCGAGCCTCATCTTTTGGCAGCAGCAGCAGCGGGGGCCAGCCGCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGAGGGGGGGTGTTACTGAACCATGTGACAGATGACGGAAGGCTCTGGAGGGCATTA TCAGCAACTTTGGGCAGACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACTCGGCTCTCA GCTGAGGAAGCAGCCCATCGCCTTGCAGGCTGCTGAAGGAGCCACATCCAACTCGGCTCTCA GCTGAGGAAGCAGCCCATCGCCTTGCACGCCTGGACACTAACTCACCTAGCATCTTCCAGCA CCTGGACGAACTCAAGGCATTCTTCGCAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGCGGTGGTGT GAGTGGACAAGACCCTGGCAGTGGCCCCGGATGGAAAGCTGCTATTCAGCGGTGGCCACTGGG ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC CTTGATGTAGCATGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCGGGCACCACCACGTGCATGGCGCTCCTGCATCAGGGTGTCTGTAGGCCTGCACCAAAGCCTGTGCAGGTCCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACTGACATGACATGGCTGTGTGTCTGGATCTGAGGATCTGACACTGTACCCCGGACCTGACACTGTGACACTGTGACACTGTGACACTGTGACACTGTGACACTGTGACACTGTACCCCG CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCCTGGGGCC CAGGTCACCTACTCCTTGCACCTGTATTCAGTCAATGGGAAGTTGCGGGCTTCACTGCCCCT GGCAGAGCAGCCTGACGGTGACAGAGGGAGTTGCGGGCTTCACTGCCCCTGGCAGAGGAGCAGCCGCCCAGTGCGCCCTGCACATCCTCCAACTAAACACACTGCTCCCGGCCGCCCCCCTTGCCCATGAAGGTGGCCATCCCCGCAGCCGCGCCTCCCTTGCCCATGAAGGTGGCCATCCCTGCAGGAGCAACCAAGGAGCGCACCACGTGCTGGTGGGCCTGGAGGA TGGCAAGCTCATCGTGGTGGTCGCGGGGCAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGC GGAAGCTGTGGCGGTCCTCGCGGCGCATCTCCCAGGTGTCCTCGGGAGAGACGGAATACAAC CCTACTGAGGCGCGC**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCCGGCAGGCCTGGCCCGGGAGGCCCCGGGAGGCCCAGGGGGTGA GCGGGGCCCACCCTGCCCAGCTCAGGGATTGGCGGCGATGTTACCCCCTCAGGGATTGGCG GGCGGAAGTCCCGCCCTCGCCGGCTGAGGGCCCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN ${\tt HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAELETP}$ MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP LAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELIP EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW IDLIFGYKQRGPAAEEALNVFYYCTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP HPTRLSAEEAAHRLARLDTNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWDGSLRVTALPRGKLL SQLSCHLDVVTCLALDTCGIYLISGSRDTTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAAVS CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGOIVVOSSA WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS GETEYNPTEAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697, 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840, 873-879, 912-918, 954-960

FIGURE 24

CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC TGGACCCTTAACTGGGTACTGGCCCTGGGCCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC GCACACTCCGTTACCACACTGGGTCATTGGCATTTTGGAGCCCTCATCCTGACCCTTGTGCAG ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC $\tt CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC$ TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA $\verb|CCTGCTGCTGTTTTTGGGAAGCTGCTGGTGGTGGTGGGGGGTCCTGTCTTTT| \\$ TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT TTTCGGCATGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC $\mathsf{GAGGCGCCCCGGACAACAAGAAGAAGAAGAAGAAGTGA}$ $\mathsf{CAGCTCCGGCCCTGATCCAGGACTGC}$ ACCCCACCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC TCGGGAGGCTGAGGCAGGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA AAAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKN
EAPPDNKKRKK

FIGURE 26

GAGTCTTGACCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT GGCT**ATG**TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT CAGGAAATGACAGTGATGGGTCAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA GTGGAGCAAACCATGCGGAGGAGGCCAGCGGCGAGAGTGGGAGGGCCCGGAGAAGAGACATCCT CTTTGACTACGAGCAGTATGAATATCATGGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCCTGCAGCG CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA GCGGCTCCAGGAGTTCCTTGCAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA TCTGGCCAGCGACGTGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC CTTTGCACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA AGTCCTTTGTGTGTTCGACAAAGAACCGGCGCTGCAAACTGCTGCCCCTGGTGATGGCTGCC CCCCTGAGCATGGAGCACGGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT $\tt CTGGACGCACTTATTTCCCTCCTGTCC{\color{red}{\textbf{TAG}}} GAATTTGATTCTTCCAGAATGACCTTCTTATT$ TATGTAACTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT TTTTATTAAATAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAYEDIFRDEEEDEEHSGNDSDGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEYHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSLCNTSYTAARFKLWSVHGQKR
LQEFLADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTFSIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLSRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFL
DALISLLS

FIGURE 28

FIGURE 29

CAGGAACCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC CACGAAGGGACGATACCAGAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG GAGTAAAGTACGCTCCGGTCACC**ATG**GTGACAGCCGCCCTGGGTCCCGTCTGGGCAGCGCTC $\mathtt{CTGCTCTTTCTCCTGATGTGAGAT}\mathtt{CCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC}$ CAGCGGCTGCCAACGGTGCTGTGACTCTGAGGACCCCCTGGATCCTGCCCATGTATCCTCAG CCTCTTCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCG GCGCCCGTGCCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCCGCAAGACGGCCCTGCACAGCGGCGAGGACTTCCAGACGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA CATGGCGACCGGCCAGTTTGCTGCTCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCATC CTGTACGCGCAGCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGCCTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA ACGACTTCGACACCTACATCACCTTCAGCGCCAGCGAGAACGCCATCTACAGCA ACGACTTCGACACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACTCAG CCTCTGGGCCACCCTCCCGGCTGGAGAGCTCAGGTGCTGGTCCCCTGCAGGGCTCAG TTTGCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTTCTGGGGAGA TGGACCTATTTTAAGAAGCTTGCTAACCTAAATATTCTAGAACTTTCCCAGCCTCGTAGCCC AGCACTTCTCAAACTTGGAAATGCATGCGAATCACCCGGGGTTCGTGTTAAATGCAGATTCT GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTCCTGGGTGATGCTG ATGGGGTCAGTCTATGAACCACACTGGAGCAACCAGGTTCTAGGACTTTCTCAATATTCTAG GCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCCAATTCCAACATTCTTAAATTCTCTCAT CCCTCCAGGGCTCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTTCTCTTGCTCAGGCC GCCTGTGTTCTGGGTGTTCAGGTGCTGCTGGTCCTCCATTACCCACTGCTCCCCAAGGCTGG TGGGACGGGGTCCCGGTGGCAGGGCAGGTATCTCCTTCCCGTTCCTCATCCACCTGCCCAG TGCTCATCGTTACAGCAAACCCCAGGGGGCCTTGGCCAGGTCAAGGGTTCTGTGAGGAGAGG ACCCAGGAGTGTGGGGGCATTTGGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA TAGCTCTCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCCCTCCTGCCTCCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT CTCTGCCCTCATGCCCTCTCACCGGCCCAGTGCCCCGACTCTCCAGGCTTTATCAAGGTGCTAAGGCCCGGGTGGGCAGCTCTCAGAGCCCTCCTCCGGCCTGGTGCTTTACCAAACACCTGCAGGAAAGGGCCACGGAAGCCCCAGGCTTTAGAGCCCTCAGCAGGTCTGGGGAGCTAGAGCAAAGGAGGACCTCAGGCCTTCCGTTTCTTCCAGGGTGGGGTGGCCTGGT

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSERSIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCGACCCGCCAGGAAAGACTG CTCCAGGGTCCCTCTGCTGCTGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCAGAGTCTTCTGCACTGCCCGCCAGGGG ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT AGAACCAGATCGCCAGCCTGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT $\tt CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC$ TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC $\tt CTGCCCGCCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT$ CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGTCTGGGGCTGCAGCAGCTGGACG AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCCCTGCGGCTGGCCGGCAACAC CCGCATTGCCCAGCTGCGGCCGAGGACCTGGCCGGCCTGGCTGCCAGGAGCTGGATG TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG CTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCTG GGTGCGCGAGAGCCACGTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAGCCACCACC ACCACAGCCACAGTGCCCACCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCTTAG CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGCCCACTGAGGCCCCCCAGCCCCCCCA CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCAGGACTGCCCACCGTCCACCTGCCTC AATGGGGGCACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTT CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGGACACGGCCCAGCCCTACACCAGTCA CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGG CCCGGGCGGTGCCGAGGGCGAGGGCCTGCGGGGGGGGCCCATACACCCCCAGCCGTCCA CTCCAACCACGCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCG GGGCGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCCTGGGCCCCT GGAACTGGAGGGAGTGAAGGTCCCCTTGGAGCCAGGCCCGAAGGCAACAGAGGCGGTGGAG AGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCCAGGGCCTGGCCTC ${\tt CAGTCACCCTCCACGCAAAGCCCTACATC} {\color{red}{\textbf{TAA}}} {\tt GCCAGAGAGAGACAGGGCAGCTGGGGCCG}$ GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCCAGAAC CGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCCTGCTGGGCTCTCCCACTCCAGGCGGA CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGGAAAGGAAGATGCTTTA GGAACATGTTTTGCTTTTTTAAAATATATATATTTTATAAGAGATCCTTTCCCATTTATTCTG GGAAGATGTTTTCAAACTCAGAGACAAGGACTTTGGTTTTTGTAAGACAAACGATGATATG AAGGCCTTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVPLLLPLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFEN GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP GILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAG NTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG PWVRESHVTLASPEETRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALS SSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLRL TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPA VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPGAG PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTTGAGCTCAT CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**AGACTCATCAG TGACACCATGCCTATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA GTGGGGCAAAAATACAAAAATCAGATTTCCAGAAAAATTGCTCATCTGCATCTAAATACTGTC TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC AAAACTGCACAATGTTTTACCAATGGACAAAATTTCTGGGTTCTTTTGCGTGATGGAATCA AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATTCCAATATTTAAATTTTGCC AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACTGCCTCACTTGAAAACTCT CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC CCTTGGAACACTTGGATCTAAAATCTATTACAACATAAAAATGATGAAAATTGCTCA TGGCCAGAAACTGTGGTCAATATGAATCTGTCATACAATAAATTGTCTGATTCTGTCTTCAGGTGCTTGCCCAAAAGTATTCAAATACCTAAATAATAACCAAATCCAAACTGTACCTA AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT CTCCCTGGATGCAGTCATTTCAGTAGACTTTCAGTTCTGAACATTGAAATGAACTTCATTCTCAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAAACTCTAAATGCGGGAAGAA ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAACTAG GTTAAAAGACGTTCATCTCCACGAATTATCTTGCAACACACGCTCTGTTGATTGTCACCATTG TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC TGGTATCTCAGGATGCTAAGTCAAACACACAAACATGGCACAGGGTTAGGAAAACAACCCA CTTTATGAAAGCTACTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA GGAAAAAAAGCATACTTGGAATGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT ATAAAATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATACACCTTTATGATGCAATTTGATACACCTTTATGATGCAATTTTATTAAAATAAAAAATGGTTATTCCCTTCATATCAGTTTCTAGAAGGATTTCTAAGAATGTTATAAGGGCTTTCTAGAAGATTTTATAAGGGCTTTCACAAGTTTATAAGGGCTTTATAAGGGCTTTATAAGGGCTTATGGAAAAAATGTGGCCAGGTTGCTTCATCCAGAATGTTTATAATCATGAAAAATGTGGCCAGGTTGCACCAGGATGCTTATAATCATGAAAAATGTGGCCAGGTTGC GAGATGGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTA GCTGGGCGTGATGGTGCACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTCCAGCCTGGT CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCCTGGTATGGAAAAAACATATTAATAT GTTATAAACTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTTGCCATTGAAATGGCATTGAA ATAAAAGTGTAAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTTGGA TTACAGGGAGCATTTGATTTCTATGTTGTATTTCTATAATGTTTGAATGT ATCTGTATTTCTTTTATAAGTAGAAAAAAAATAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQLDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNDENCSWPETVVNMNLSYNKLS
DSVFRCLPKSIQILDLNNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVLGLAVAFCCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHHNLFHENSDHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTTGGAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTCGCCGGGAAAGG GAGGGAAGAAGGAAGGCCGGCCCGCCCCTGCGCCCCGCGCCCCTGTCCGCCCCCGGC CTGGCAGTGACCCTGGCCGGGGTCGGAGCCCAGGGCCCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT AAGAGGGAGAAGTCGGCTCCGGAGCCGCCTCCACCAGGTAAACACAGCAACAAAAAGTTATGAGAACCAAGAG TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA CATCGAGGGAGCTCAACATCCAGGCGGGCATTAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAG AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCATCACTCAAGGGA GGAACTCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTC ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGA GAATGGAGATCCTGGGCTGCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT GATGACCTGGATTTAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAAATGTGTCC CAATATCACCAGAATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCCACGGCAATGAGGTGCTGGGC $\tt CGGGAGCTGCTGCTGCTGCTGCTGTGTCAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT$ GGTGGAGGAGCGCGGATTCACGTCCTCCCTCCCTCAACCCCGATGGCTACGAGAGGCCTACGAAGGGGGGCT CGGAGCTGGGAGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACTTTCCTGATTTA AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCCAATCACTATATTGCAATCCC TGGAAGACGCAGGAACACCCCCCACCCCGATGACCACGTGTTCCGCTGGCCTACTCCTATGCCTCCAC ACACCGCCTCATGACAGACGCCCGGAGGAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGGCACTGTCA ATGGGGCCTCCTGGCACCCGTCGCTGGAAGTCTGAACGATTTCAGCTACCTTCATACAAACTGCTTCGAACTG TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATC TCTGATCGTGTTCATGGAGCAGGTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTCACATGGAAAAGGAATCC CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGGATTACTGGCCCCTC $\tt CTGAACCCTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG$ CTATGACATGGGGGCCACAAGGTGTGACTTCACACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGG TGGACTCACTCTTGTTTCCTCTGTAATTCAAGAAGTGCCTGGAAGAGGGTGCATTGTGAGGCAGGTCC CAAAAGGGAAGGCTGGAGGCTGTTTTCTTTTCTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCA ${\tt GAGCCTGTCCGTTCAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGCGTGGCAGCAAGG}$ GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAAATGTTACCA TTTGAGATGCTCCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA ${ t ATTCTGTGTTTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT$ ${\tt TGGAGCTTCTTGCACAAATTCTGGGTCCATAAACAACCCCCAAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT}$ CCCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC GAATTGAGTGCTCATGGGCTTGGCCTCATATCAGCCTGGGAGTTATTTTTGATATGTAGAATGCCAGATCTTCCA GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGAALEDPDYYGQEIWSREPYYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSYAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSKTNMARIREIMEKFGKQPVSLPARR
LKLRGRKRRQRG

FIGURE 37

ATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC TTCAGCTCCAGGTCGGGCTCCAGCTCCAGCCGCAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT TTCCAATTTCACCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAATTGACATCAT GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT ATGACTCTCTTGGTAGAGAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTCGCCGAGAAATCGTGGC ${ t TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCTCCCCACTC}$ ${ t CAGGGAGCTGTGGTGGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGGTTT}$ ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACTCTCCCTAA TGCTGCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGGTTCT GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGG GCAAACTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACTATAACCCTTTTGACCAGAAA $\tt CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAG{\color{red}{\textbf{TAA}}} GCTGTTTA$ $\tt CTAAAAGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT$ TTGTCAGAGGTCTAGGGGCCACTGTGGGCCCTAGTGAAGCCTTACTGTGAGGAGGCCTTCACTAGAAGCCTTAAATTA GGAATTAAGGAACTTAAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC GGAGCTCCTCGAGGGACCAAATCTCCAACTTTTTTTTCCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG ${\tt CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT}$ ATGCATTAAATTGTACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTTCTCTC AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACTTGTAAGAAAAT GAAAATTTAATTTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT

FIGURE 38

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

FIGURE 40

CCGCTGCTCTTGTGACGTTGTGGAGATGGGGAGCGTCCTGGGGCTGTGCTCCATGGCGAGCT GGATACCATGTTTGTGTGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAAC GGTTTGGCTATGTTCTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA TTATTGGGGCATTCTTCATTCCAGAAGGAACTTTTACAACTGTGTGGTTTTATGTAGGCATG TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC TACACTCATCCAGCCAGTTGTTCAGAAAACAAGGCGTTCATCAGTGTCAACATGCTCCTCTG TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC TGTCCCAAAGGAAGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC TCTTTTTGTTGTGTGTATTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA CTGACTCTAACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC ACTGGAGGATGGGGACGATGTTCACCGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA GTTATTCCTTCTTTCACTTCATGCTTTTCCTGGCTTCACTTTATATCATGATGACCCTTACC AACTGGTCCAGGTATGAACCCTCTCGTGAGATGAAAAGTCAGTGGACAGCTGTCTGGGTGAA AATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCACTTGTTC $\tt TTACAAATCGTGATTTTGAC \underline{\textbf{TGA}} \texttt{GTGAGACTTCTAGCATGAAAGTCCCACTTTGATTATTGC}$ $\mathtt{TTATTTGAAAACAGTATTCC}\overline{\mathtt{CAA}}\mathtt{CTTTTGTAAAGTTGTGTATGTTTTTGCTTCCCATGTAAC$ TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTCATTTTGTTATTTTCTTACCAA GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTTATGAATATGGTGATGAGT ATGGTTTATTTTAAAATTTATAAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG TTATGGATGGAGGTGTCGGTACTAAATTGAATAACGAGTAAATAATCTTACTTGGGTAGAGA TGGCCTTTGCCAACAAAGTGAACTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTCAGT GGAAATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA AGGAAGTGTTTTGAAAGTCACTTTGAAAGTTAGTTTTGGGCCCAGCACGGTAGCTCACCCTT GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA GCTTGGCACATGGTGAACCTGTTCTATAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCCAGAGCCAAAGGTTGCAGTGAGCAAGTCA ATAAAATCTAGTCCAGTTCTCTCATTTAAAAAAATGAAGACACTGAAATACAGACTTAAATA GATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCAGAAGAATGATGGAATGTTTT AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTAAAATGCTTTTGTAT TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA ATGAATTCAGAGAAAAAAAAAAAAAAAA

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVWVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAANTCCACTGTAACTAGATTGATCTA
TGCACTTTTCTTGCTTGGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

FIGURE 44

FIGURE 45

FIGURE 46

GGCAAATGTGTGTGGCTGGAGGCGAGCCGAGGCTTTCGGCAAAGGCAGTCGAGTGTTTGCAGACCGGGGCGAG GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC $\verb|CCCTGGTTTGTGTGTTACGCACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCGTTTCCAGCTCC|\\$ AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGAT TGGAAGCTCGGCCTTCTGTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCA ACATCATCCTGGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC AGGCACAGCACGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTCGGACTCCTTAAAAACTCCCG CTTTTATAACTACACGCTGTGTCGGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA CAGACCTCATCACCAATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCCTC ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCTCTTCCCAAACGC ATCTCAGCACATCACGCCGAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC TCCCGTTCTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCCTCAACATTGACCTG GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG AGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTAC CAGCGTGTGAAGGACCTGTGTCAGCGTGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC AGCCTGGCCGGACGCCGGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG $\verb|CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCCAGCCCCGAAACCTCA| \\$ CCAAGCGGCACTGGCCAGGGGCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC CTTCCCGACTACTCAGCCGCCAACCCCATTAAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA CCCTGCAGAACAAATTAAGAACCTGAGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCCAGAAGAATGTGAC TGTCACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAG GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAAAACTCCGCAAGCTGC TCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG ${\tt CAGACGGCGCCTTTCTGGGGCCTTTCTGTGCCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT}$ GAGGACCATCAATGAGACTCACAATTTCCTCTTCTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCA ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACACGTACAG AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG ${\tt GACAACTGTGGGGAAGGTTAA}{\tt GAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA}$ $\tt CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC$ AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC ATTTTTGCCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC ${\tt TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTTATTCTT}$ TTGGTTTGTCACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAAACAGTGCAGAGACG TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCCTGGTTGCCTCTGAAGAAA CTGCCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAAT CCCAATTTTCAGGAGTGGTGGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHES
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAG**ATG**CTGGGCCTCCTGGGGAGCACAGCCC TCGTGGGATGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCCACC TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGCAGCCGTACAGCTGCAGGGGGAAA CCGAGTCCGCCGGCCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC CCCCGCCACACCCTCACCACCTCCACCACCACCACCCCCACCGCCACCATCCCCGCCA CGCTCGCTGAGGCTGCCGGTGCCTGTGGACAGCAGCTGCCCTGCCCTCCCATCTG TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG ACAAGCGTGTCTTGGCAGAGCCAGCACACAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC ATCAGGCTGCTGCAGGCCTCTGGCGGCAGGGCACTGGGAAGGCCCTGAGAATGTCCTTTT GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT

FIGURE 49

 ${\tt MLGLLGSTALVGWITGAAVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRAQPWPFR} \\ {\tt RRGHLGIFHHRRPGHVSHVPNVGLHHHHPRHTPHHLHHHHPRHAR} \\$

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCCATGTCGGACCTGCTA CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG GTACTCAGGGCTACTGGCTGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCA $\tt CTGTGGCCTACAAGTTCCACATGGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC$ TGCAGCATCTCCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG AGATCTACCAGGAAGACCAGATCCATTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT GTGCCTGAGATGAAGGAGACAGAGTGGAAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA GGAGCTGGACTTGGAGGGCCAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAGT**TAA**CCC ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT $\tt CCAGCCCTCTTCCTCCTTCCTCTGGGGGAGGGGGTTCCTGAGGGACCTGACTTCCCCTGC$ TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA GGGACTATTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTCAGACTC ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAA AAAAAAAAA

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPLAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGPLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

FIGURE 52

FIGURE 53

MTLRPSLLPLHLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS KKKLKEEKRNKSKKK

FIGURE 54

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAAACGAAA
CAAGAGCAAAAAGAAATAATAAATAATAAATTTTAAAAAACTTAAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG TGAGGCGGGCGCGCGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGCTGGACTG GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT GTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT ${\tt TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAG{\tt TGA}{\tt A}CACATCTGAT}$ $\tt TTCCCACAGCACAACAGCCCTGCATGGGTTTGTTTGTTTTTTTACTGCTCACTCCCAACCTT$ TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT AAAATCACGAGAACACCTAAACAACCAAAAAATCTATTGTGGTATGCACTTGATTAACTT ATAAAATGTTAGAGGAAACTTTCACATGAATAATTTTTTGTCAAATTTTATCATGGTATAATT TGTAAAAATAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAAATATTCCGTGG TCAAAATTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT GTAACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT $\tt CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC$ CACATCCACCACTG

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWWIIIDAAVIYPTMKDFNHSYHACGVI ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

FIGURE 58

FIGURE 59

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTCCAGAATGCC

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

FIGURE 62

FIGURE 63

CGACGCCGGCGTG**ATG**TGGCTTCCGCTGGTGCTGCTGGCTGCTGCTGCCGTCC TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC AAACGGCCCCCAGCGCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC TTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGCTGGAACAACAT CTTACGGTGTCACCCCCAACCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC ATGAAAGGAGCTTTTATCCCCGAGGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT GATTCAGCGGGCTGGGGGCGCTGTCCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT CAGCTGGGAAAGCCTGTGGTGTCAGTGTAAGAAGGGGCATGAGCTGGTGAACATCTATTGC CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC CCGCTGCCTGCCAGGTGTAAGCACCTATGAGCACCTACTGCCGGCCCGGCATGACCCT CTTTGTGGAAGCCTCTATGTCAGTGGTCCTGAAACTGTTCCCACAGCTGGAGGGGAAGGTGG CTTTGTGGAAGCCTCTATGTCAGTGGTCCTGAAACTGTTCCCACAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTCACCAACCAGTTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGCCGCCTGCACCCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACCTGTGGACTGGTCG
GGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACTTGTACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAATTAGTTCCATCAGGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAAGTTCCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAATTCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTACACTCTTTATAACATGCCATCCCTACTAATATGACTTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTTCACTCAACCTGGTGGGTTCAAGTG
TGTCCTGAGGCTTCTTCATTCATTTAGTGCTACGCTCCACCTGGTGGGTTCAAGCTC
TGTCCTGAGGCTTCTGCTCTCATTCATTTAGTGCTACGCTTGCACGTTCTACACTGGTTCAAGCT CTTATCCACCAAATACACAGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG CACAGCAGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAAGTCACATCAGAAAAGGGA AAGCCACGGAATGTGTGAAGCCCAGAAATGGCATTTGCAGTTAATTAGCACATGTGAGGG TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTG GTATCAGACATACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTG ATTCCATTGCTTTAAAAAAAAAAAAAAA

FIGURE 64

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ EEAIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMPREEAAEH IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEEWQAELKGKRGSDYETFKNSFVEA SMSVVLKLFPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

FIGURE 65

GCAGCGGCGAGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA TGCGGTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT GTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGAATTAGAATCCTCTATTCAAGA AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGGAAAGTGTCACAGAAGATATCAGCTTTCTAG ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCCTAGATAA GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTTGTGAAACTGAAGAAGAGGCTGCTAAGAGA CGGCAGATGCAGGAAGCAGAATGATGTATCAAACTGGAATGAAAATCCTTAATGGAAGCAA CCAAAGCCCTGGAGAGAGTGTCATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC CAGGCAGCGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTAT ATTATACATTTGGAGCTCTTGGGGGCCAATCTAATAGCCCCACATGGTTTTGGTAAGTAGACTT **TAG**TGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTT ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA ATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGC TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA TTTTAAAAATTTTTCTTTAGTTTTTCAAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAAATTCATGGAGTTATTT CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACTCATTTTTAATAAA

FIGURE 66

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSKTTLTSDESVKDHTTAGRVVAGQIFLD SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG EPCHFPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAAKRRQMQEAEMM YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEK LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMVLVSRL

FIGURE 67

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAG SAPRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVQPEDDADYYCSVGYGFSP

FIGURE 69

GCCGCCCGGGGGCCGGGGCCGGGGCGGGGCGGGATGCGGCGCCCGGGGCGCG CCCGCCGAGGTCCGGACAGGCCGAGATGACGCCGAGCCCCTGTTGCTGCTGCTGCCGC CGCTGCTGCTGGGGGCCTTCCCACCGGCCGCCGCCCGAGGCCCCCCAAAGATGGCGGAC AAGGTGGTCCCACGGCAGGTGGCCCGGCTGGGCCGCACTGTGCGGCTGCAGTGGAGGGGGCCCGCCGCCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA GCGCTTCCGCGTGCCGCCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG GGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACTCCACCATCGATGTGGGCGGCCAGAAGTTTGTGGTGCCCACGGGTGACGTGTGGTCGCGGCCCGACGCTCCTACCTCAATAAGCTGCTCATCACCCGTGCCCCAGGACGATGTACATCTGCCTTGGCGCCAACACCATGGGCTACAGCTTCCTCACCGTGCCCAGACCCCAGGCCCCAGGCCCTTCCTCACCGTGCTGCCAGACCCCAAAACCGCCAGGGCC ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCTACGCACCGT TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACACGCAGACTGACGTGCTTTTGG GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTTCCCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC GTCCAGGCCAGACACCACCCCCACCCCACTGTCGTGGTGGCCCCAGATCTCTGTAATTTTA

FIGURE 70

MTPSPLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGK
ESLGPDSSSGGQEDPASQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTHSHVEGKV
HQHIHYQC

FIGURE 71

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAAACTTCCCAGGGGACCGCATTCCAGAGTC CCAGTGTAGCCCAGGGCAGATGTCAATAAATGCATACTCTGTATTTCGAAAAAA

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGROTMLTOSVRRVOPGKKNPSIFAKPADTLESPGEWTTWFNI DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGOVVHGSPREGFWCLNREO RPGQNCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTK TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPY MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQNATNSFYYDVGRCPV KTCAGQQDNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRV SAADNGEPMRFGHVYMGNSRVSMTGYKGTFTLHVPQDTERLVLTFVDRLQKFVNTTKVLPFN KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRONGEPYIGKV KASVTFLDPRNISTATAAQTDLNFINDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL DVPESRRCFVKVRAYRSERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPOPYLNKLNYRRTDHEDPR VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN EDDPMSWTEDYLAWWPKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTLVKVIPQGSCRRASVNPMLHEYLVNHLPLAV NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA QQPLIN

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC ACCTACCTACCCGTACGCATACATACATATGTGTATATATGTAAACTAGACAAAGATCGCAGATCATAAAGC AAGCTCTGCTTTAGTTTCCAAGAAGATTACAAAGAATTTAGAGATGTATTTGTCAAGATCCCTGTCGATTCATG CCCTTTGGGTTACGGTGTCCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCTGAGCTGATGTTTGATTTT GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA GATGCTTTTCACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTCGCGCTTTTTG CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTGTATGTG TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA ACAACGTGCGCTGCCTGTGCCCGGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGC AGCTGCGGCTCCGACTCTGGCCAGGGCGCGCCCCGCACGGCACCCCAGCGCTGCTGCTGACCACGCTGCT GGGAACCGCCAGCCCCTGGTGTTC<u>TAG</u>GTGTCACCTCCAGCCACACCGGACGGGCCTGTGCCGTGGGGAAGCA CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTC TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC TGTGTAACAGCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA ATTTTTCTTGAACTACTGTAATATGTAGATTTTTTGTATTACTGCCAATTTGTGTTACCAGACAATCTGTTAAT GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATATAACAAGTGTAATAAGA TTCCACCAAAGGACATTCTAAATGTTTTCTTGTTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA TTACTGATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCCAGGAAGATGGATTGATGTTCACT AAAAAAA

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD TTKKLRDFFTVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS PLVF

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGG CGGGTAGCTGGCAGGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCCTCATGACTT CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC CCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTTTTTTGGAATGGAGGAGACGGCTCA AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT AGTATACATTTTTCTTCTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAG AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAAAAAAT TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT CATTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT TGGACCAAACCTCG**TGA**GCCAGCCACCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT CCGGGAGCAGTGATGTCAAACTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCGGAAACATCTGCCCATGTGTATTG ATGGCAGAGCTGTTGCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT TCCTCTGACCCATGCTTAGTACATATGACCCTTTAACCCTTACATTTATATGATTCTGGGGTT GCTTCAGAAGTGTTATTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT TTACTTTACATATTCGTTTTCAATACTTGCTGTTCATGTTACACAAGCTTCTTACGGTTTTC TTGTAACAATAATATTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC TAAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG **AAAAG**

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVKW HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG CCTGCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT TCTGTCGCAGGCTGCGAGGAAAGGCCCCTAGGCTGGGTCTTGGCGGCGGCGGCGCTT $\verb|CCTCCCGGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGT|| \textbf{A} \\$ **TG**GAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCCCC TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACTACTACATCCAGTGGCT CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTGTCCCTCA TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG GGTGTCCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT AGGTATGGTGTGGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT CTGCTCCTGGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGCGGAAGGC GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG $\verb|CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG|\\$ TGCCGTCATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCCTCAGTTGTGGGCCTTCTATA GCTCTCCACTCTTCCGGAGCCTGCGGCCCAGATGGCACGACACTGCCATGACGCAGATAATT GGGAACTGTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG ACTGCAGCTGTGCGGCCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAGTGACCTCCAGCTGGGGGTGGGA AGGAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG AAAGACCAAGGGGATAGGGAGGAGGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA AAAAA

FIGURE 79

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDEDATVNK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGVLGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMPRGMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVGFY
SSPLFRSLRPRWHDTAMTQIIGNCVCLLVLSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY
IVFLYNAAFAGLTTLCLVKTFTAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

FIGURE 80

GGCTGCCGAGGGAAGGCCCCTTGGGTTGGTCTTGGTTGCTTGGCGGCGGCGGTTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

FIGURE 81

FIGURE 82

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCGGTGT GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG ${\tt GAGTGTGTGGAACAGGACCCGGGACAGAGGGAACCT} {\tt GAGTGTGTGGGAACCTGAGCACCTTTT}$ GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCA GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG CTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTCACACTTCTTTGGGGATTT TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGAT GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT GCCCTAATGTCAAACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG AGAGACGCCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA ${\tt CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT} {\color{red}{\bf TGA}} {\tt GAGTG}$ TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACC AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA GTTGTTAGCAATTTCATTCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTG TTATTTTTA

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRKQYDTYGEEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPRQQ
DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFIGEGEPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGAGAGGAGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATG**G AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG TACCCTCCGTTGGACCCCAAACTCCTGGACGCCCGGACGACTGCCCTGCTCTGTCAG TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC AGTCTCTGTCGGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG ${\tt CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATT} {\color{red} {\bf TAG} {\tt TGCCT} }$ ACAGGCCAGCAGCTAGCCATGAAGGCCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACTGTGGCTGGTGAGT GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGGAGTATTGAAAA CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAGGTTTGGGTTTGAAGCTGAGGAACT ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWI DQSLSAAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCC CCACCACTGCAGCC<u>ATG</u>ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA TTTGGAGTGTTTTTCCTGTTCTTTTGGAATGATTCTCTTTTTTTGACAAAGCACTACTGGCTAT TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTC CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTCAG GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTATAAAGTCATTTGAAGAATATTCA GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA ACTAAGAAGTCAGCAAGCAAACTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA ACTCTTGAAGGCTATTTGTGTTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA $\tt CTGTGGTGCCTGTTTCTTTTTTTTTTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT$ TTTTAGAAGTGTCCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA TGCATGAATTCGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIRRVPVLGSLLNLPGI RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

FIGURE 91

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA CTGTGGGTAGTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC ATCAACACCATTCAGCTCTTCACTCTCCTCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCCT GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC GGGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAA TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC AAGCTCTACCAGGAGAAGGATGCCTTTCAGGAGGAGTACTACAGGACGGCCACCTTCCCAGA GACGCCCATGGTGCCCCCCGGCGGCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGC TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAGCAGAAACTGAATGAC**T GACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACTGGTGGCCTCTGCATATCCT** CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIASGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQHLRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

FIGURE 94

 $\tt CTGAGGCGGGTAGC \underline{\textbf{ATG}} GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTTGTGCTCG$ GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT AAACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTTATAAAAC TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT AGTAAAGGATGTAAAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG CAGCAAGAGAAGAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACA TGTTTCTAAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT AAAGCCTTAGACTAGATGACAGATGCCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC CAGAAACAGATGAAGAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA TTT**TGA**TCCTTTTAACCTTACAAGGAGATTTTTTTTTTTTTGGCTGATGGGTAAAGCCAAACAT TTCTATTGTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCAC CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC ATCAGATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCT TACACAGACACTTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDTQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

FIGURE 96

GGCACAGCCGCGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC CCTTCCCCGCGTCCCCGCCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCCCCCCGCGAAACCCCGAGGTCACCA GCCCGCGCCTCTGCTTCCCTGGGCCGCGCCGCCTCCACGCCCTCCTTCTCCCCTGGCCCGGCGCCTGGCACC GGGGACCGTTGCCTGACGCGAGGCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCCTGCTCGCCTCTTCCAC CAACTCCAACTCCTCCCCCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAG $\tt CGCCGCTTCCCGTCCGAAAGGTGGGAACGCGTCCGCCCCGGCCCGCACC \underline{\textbf{ATG}} \texttt{GCACGGTTCGGCTTGCC}$ AAGTGCGACGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCCTCCACGAGATCAACGGTGATCAT TTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA TGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTG ATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT TTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAACTCCCAGT ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT AGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCCACAGCCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT GCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACTACTGCTCAAACATCATGAGAGGCTGTTTG GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA GGGTCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG ATAATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCCTCCCAGCTGGACGAATT TCTCGTTCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGC AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGTCCT CCCTTCCGAGCAACGTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA GAAAGCCGACAGTGCTGGTCCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGG $\verb|TTATGCAGAGAGAGTGGAGA| \verb|TAA| \verb|TTCTCAAACTCTGAGAAAAAGTGTTCATCAAAAAGTTAAAAGGCACCAGTT| \\$ CACTGGTTTAAGAAGTGCTGACTTTGTTTTCTCATTCAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT CGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAAGAAAAAGCCCAAAAAGC

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLYMQNSELFKDLFVELKRYYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEY
LECVSKYTEQLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGCGPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLF
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

FIGURE 99

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR KFMTVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

FIGURE 100

FIGURE 101

MAVLVLRLTVVLGLLVLFLTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVE FILRSMSRSTGFMEFDDNEGKHSSK

FIGURE 102

GGACGCCAGCGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT ${\tt CAGAGCTGGTCTGCC} \underline{\textbf{ATG}} \texttt{GACATCCTGGTCCCACTCCTGCAGCTGCTGCTGCTTCTTAC}$ CCTGCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCC CCTACCTGATGGCCGTGCTGACTCCCAAGAGCCAACCGCAAGATGGAGAGCAAGAACGGGAG CTCTTCAGCCAGATAAAGGGGCCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA AGCTGGGCCTTCATGTGGCAGCAAGTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG CTGCCTCACCAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG AACGACAGCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACC**TAG**CAGAATGAGAGAGACATT CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT CTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT **AACCACG**

FIGURE 103

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQI KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV APGEDMRQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFM WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP SSKALICSFPSLQLEQATHQPIYLPLRGT

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAGTTGCCTCATCGCAGG CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA ACTAATATTTATATGACAGAAGAAAAAG**ATG**TCATTCCGTAAAGTAAACATCATCTTGG TCCTGGCTGTTGCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAAA TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGGAGATTCCTGTGGTCATCGCTGCATCTG AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTCAGCACAACACTCGCTCCAAT GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAG AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGTAATTGTGCA AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG AAGATTGTGATTCAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG CTCATTTAATCCTGGAGTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGGGGACTGTATAGCAGAACCCTGGCT GGTAGCATCACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTAA AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAG ATATACCGAGATCTCAAACATAAAG**TGA**AACAGAATTTGAACTGTAAGCAAGCATTTCTCAG GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAAACTGCCCTGTCTGGCAGTC AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT ACATTTTTC

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR QEEIPVVIAASEDRLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA NLTEWKRQNITNQLEKWMKLNVEEGLYSRTLAGSITTPPLLIVFYQQHSTIDPMWNVRHLGS SAGKRYSPQFVKAAKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

FIGURE 106

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG TCCGCATCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT TTCCGATCACTTCAAAGTGGTTAAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT CCGTAAAGTAAACATCATCTTGGTCCTGGGCTGTTGCTCTTCTTACTGGTTTTGCAC CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTCAGGAATTGTAG GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA ACAGCATTCAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG TCAATTTTGACCCTAAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG CCATATACATGGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTTGCAAAC CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGT AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCG TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTTGGGGAAAAATGGTATATTCCA GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAAACTGCTGTTGGTTTTAATTTT AAAAA

FIGURE 108

FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH PRSPAMKAATCCSPEGPWPSLEPRT

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA GTTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAA TCATGTCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC TTTGGGATTGTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCTCTGGGTGGCTGCTGCTG AGCCTGGGAACTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCT TTCGGGCATTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA TACTACAACTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT TGTCCAAGAACTCAAGTCACTTTACATCTATTAACTGCTTTGGAGACTTCATAATTTTTCTA GGAAAGGTGTTAGTGGTGTTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG ATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA $\tt ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGA\underline{TAG} ATACCCATTTAGGTATCTGTACCT$ GGAAAACATTTCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT AGTGAATTTTTTTTAAAAGACCTAATAAACCCTATTCTTCCTCAAAA

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVILGLLFVCGVLWWLYYDYTNDL SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN EEGTELQAIVR

FIGURE 112

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCCTT TGTGGTGAAAATTTTTTGAAAAAAAATTGCCTTCTTCAAACAAGGGTGTCATTCTGATATT T**ATG**AGGACTGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCCTTGTTTTTGCTGG TGACTGGAGTACATTCAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCCAAGTTCACTGTG CCTCAGATCAACTGCGATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATG TCCAGCAGGATGCCAAGACCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACT CCAGTGTGTGTGGCGCTGCCGTACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTT GTTCGGAAGGTTGCTGGACAGTGGTGCTTGATAATTCAGGAGGGAAATACTT GTTCGGAAGGTTGCTGGACAGTCTGGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTT ATCCTACCACGATGGAGAGAATCCTTTATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAA CCTACCCATCAGCTCTTACATACTCATCGAAAAGTCCAGCTGCCCAAGCAGGTGAGACC ACAAAAGCCTATCAGAGGCCACCTATTCCAGGGACAACTGCACAGCCGGTCACTCTGATGCA GCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCCACCACCTTGCCAAGGCCATCCCCTTCTG CTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGGGCCACAGGAGCCAGGAGATGGAT CTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCCCAGAGCTGATCCAGGTATCCA AAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGCGGATGTCAGCCTGGGAC TTGTTCCAAAAGAAGAATTGAGCACACAGTCTTTGGAGCCAGTATCCCTGGGAGATCCAAAC TGCAAAATTGACTTGTCGTTTTTAATTGATGGGAGCACCAGCATTGGCAAACGGCGATTCCG AATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGCCGGTCCAC AATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAATGT AGGTCGGGCCATCTCCTTTGTGACCAAGAACTTCTTTTCCAAAGCCAATGGAAACAGAAGCG GGGCTCCCAATGTGGTGGTGGTGATGGTGGATGGCCCACGGACAAAGTGGAGGAGGCT TCAAGACTTGCGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGA AAATGAGAAGCAGTATGTGGTGGAGCCCAACTTTGCAAACAAGGCCGTGTGCAGAACAAACG CAGTACACCTACGAACAGCGGCTGGAGTTTGGGTTCGACAAGTACAGCAGCAAGCCTGACAT CCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCACAGCCTCGGAAC<u>TGA</u>ATTCAGAG

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQ
RQDPSGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLQPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDI
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

FIGURE 114

 ${\tt CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCGGGGGGCGCTGAGAGGACACGAGCTCT} {\color{red} {\bf A}}$ TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT GCGCCAGGTCCCGACGGCTCCGCCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACTCC AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGC GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC AAGTGGGTGGCCCAGGACGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCCTAGAGTATAACAAAG CCATCCGGAACTACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT GTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT CCTAGAACTCGGAAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG GTCCACCTTCGACGCGGTGATCACCCCCTATGGGGGGTGCATCCTGGGGGGCTGGGGGGTACA TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG GAAGAGCAGTGGGAGGAGCTTGATGAGGGAATTCTACTCTCAAACGGAGCAGGTC GAAATTTCAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGGGAAGCCTGCCAAACAGAAGGTCCCA $\tt CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCCT$ AATTGCTTTTGGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLWVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

FIGURE 116

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG ${\tt GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA}$ GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGGCCTGTACACGAGCCACATCT GGATCCCCAGCAGCTGGTGCTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA ${\tt GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT}$ ${\tt GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG}$ TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT ${\tt CTCA} \underline{\textbf{TAG}} {\tt GTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC}$ ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAAGTAGGAAGA GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC $\tt CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC$ TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGGAATGGCTTAGCGAGCTCTACAGT AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA AAAAAAAAA

FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

FIGURE 119

 $\tt CGGACGCGTGGGCCCACCTCCGGAACAAGCC{\color{red} \underline{\textbf{ATG}}} \texttt{GTGGCGGCGACGGTGGCAGCGGCGTG}$ GCTGCTCCTGTGGGCCTGCGCCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG TCAACATCCGGGGCAAACTGGTGTCGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG AGACCTGGGCCCCCACCACTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCGCCGCACCTACAGTGTCTCATTCCCC ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA GACTTCTGGGAAGGAGCCCACCTGGAACTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG TGGTAGGGGCTTGGGACCCAACTGTGTCAGTGGAGGGGGTCAGACCCCAGATCACAGCGCTC $\tt GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA{\color{red}{\textbf{TAA}}}CCACCGCGTCTCCTCCTCCACCA$ CCTCATCCCGCCCACCTGTGTGGGGGCTGACCAATGCAAACTCAAATGGTGCTTCAAAGGGAG AGACCCACTGACTCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAATGTGTGGCAAA TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT GTTATTTCCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA AACAATACCTCACGATATAAAATAAAATGAAAGTATCCTCCTCAAAAA

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ HYRALQQLQRDLGPHHFNVLAFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG AHPAFKYLAQTSGKEPTWNFWKYLVAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

FIGURE 121

 $\tt CGGACGCGTGGGCCGGGACGCAGGGCAAAGCGAGCC{\color{red} ATG} GCTGTCTACGTCGGGATGC \\$ CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAGGGGTGCACCAAAAAGCATCTTA ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC ${\tt TTGGTCGTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA}$ AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC CTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG TCTGTGAACCCAGCCTACCAGGCTATGGAACTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT GTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCTTGAAGAGTCAGAGGCTCCCAGATCTGACC ACAGTCATCTCGGTGGATGCCCCTTTGCCGGGGACCCTGCTCCTGGATGAAGTGGTGGCGGC TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCCTGTCCTGCCATG ACCCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCCACCCTCTCC CACTACAACATTGTCAACAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAG GCACAATGATGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG GTGTCATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGG CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGC ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG CTCGAGGACTTCTTTCACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA CGATCGGATGGGGGAAGAGTTTGTGCCTGCATTCGGCTGAAGGACGGGGAGGAGCCACGG TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC GTGTTTGTCACAAACTACCCCCTCACCATTTCAGGAAAGATCCAGAAATTCAAACTTCGAGA ${\tt GCAGATGGAACGACATCTAAATCTG} \underline{{\tt TGA}}{\tt ATAAAGCAGCAGGCCTGTCCTGGCCGGTTGGCTT}$ GACTCTCTCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC TCCATCCCCACATTCCCCTGTCTGTCTTTTTTGGCATAAAGAGCTTCTGTTTTCTTT GAAAAAAAAAAAAA

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDFFHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277, 314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

FIGURE 124

GAGCAGGACGGAGCCATGATCTGGACTGC AGGCTGGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC AGTGCGGGGTTGCGGTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCT GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA GCTAATGTGACTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACT CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG ACCCACATCCACCACAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG ${\tt TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCCTACTG} \underline{{\tt T}}$ **GA**GCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT CATCACTTCCTGTTCCCACCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTTGGGAAATAAAATACCGTTGTATAT ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC TCCGCTTGTCCTCTTGTGATGTTAGGACAGAGTGAGAGAGTCAGCTGTCACGGGGAAGGTG GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC CCAATTCGCCCTATAGTGAGTCGTA

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPPLVRLPPPEPTT VASTTSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVEHEASRDEEPRLTGGAAGHQDRSNSG QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

FIGURE 126

 $\tt CGGGACTCGGCGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC{\color{red} {\bf ATG}}{\bf GAGT}$ TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGGACTGGTGT CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCT**TAA**GAAAACCGGCCACTTC AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCCTATCCCCTCTA ACACCATTCCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCCT CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTGCCCCGTGGCCCTCCATCAC CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC GTAGGGTAAGAGCACGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCAC TTTGTCATCATTCTTCATGGACTCCTTTCACTCCTTTAACAAAAACCTTGCTTCCTTATCCC ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG AGGAGTCCCCATCTGCCCCGCCCCTTCACAGAGCGCCCGGGGGATTCCAGGCCCAGGGCTTCT ACTCTGCCCCTGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG GTTGGGGCCAGCACCCGGGATGGATGGAGGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG TCCCCTTAGATGGGCAGCAGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA GCGGTGAGCGAGGTTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG GTTGAAGGTCATAACGAGAGTGGGAACTCAACCCAGATCCCGCCCCTCCTGTCCTCTGTGTT CCCGCGGAAACCAACCAAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT CCTCAACAACAACAGAAAAAGGAATAAAATATCCTTTGTTTCCT

FIGURE 127

 ${\tt MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCK} \\ {\tt CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLEPLDANVDA}$

FIGURE 128

FIGURE 129

 ${\tt MKIPVLPAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSAFKADE}$ ${\tt FLNWHALFESIKRKLPFLNWDAFPKLKGLRSATPDAQ}$

FIGURE 130

FIGURE 131

 ${\tt MGVEIAFASVILTCLSLLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS}$

FIGURE 132

 ${\tt GGGGAATCTGCAGTAGGTCTGCCGGCG} \underline{\textbf{ATG}} {\tt GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG}$ GCTGCTGTTGTTCCTCCTGCCCTCAGCGCAGGGCCCCAGAAGGAGTCAGGTTCAAAATGGA AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC TGCAGCTGCTACCATGGTGTCATAGAAGAGGGTCTAACTCCTTTCCGAGGAGGCATCTCCAG GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTTGAGCACTTTATT TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACTTGTTGATGCAGAATACACCAAAAAC CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT ${\tt TCCCAAAATGTTGAAAACTGAACTA}$ ${\tt TAG}$ ${\tt TAG}$ ${\tt TAGTCATAGGACCATAGTCCTCTTTGTGGCA}$ ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA TCTGCTATCAAGCCAAATACCTGGTTTTCCTTATCATGCTGCACCCAGAGCAACTCTTGAGA AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAAAACCATAAACTCTGTTACTCAG GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

FIGURE 133

MEWWASSPLRLWLLLFLLPSAQGRQKESGSKWKVFIDQINRSLENYEPCSSQNCSCYHGVIE EDLTPFRGGISRKMMAEVVRRKLGTHYQITKNRLYRENDCMFPSRCSGVEHFILEVIGRLPD MEMVINVRDYPQVPKWMEPAIPVFSFSKTSEYHDIMYPAWTFWEGGPAVWPIYPTGLGRWDL FREDLVRSAAQWPWKKKNSTAYFRGSRTSPERDPLILLSRKNPKLVDAEYTKNQAWKSMKDT LGKPAAKDVHLVDHCKYKYLFNFRGVAASFRFKHLFLCGSLVFHVGDEWLEFFYPQLKPWVH YIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMDDITCYWENLLSEYSKFLSY NVTRRKGYDQIIPKMLKTEL

FIGURE 134

 $\texttt{CACCCTTCCATTTCTCGCC} \underline{\textbf{ATG}} \texttt{GCCCCTGCACTGCTCCTGATCCCTGCTCCTCTT}$ ${\tt TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT}$ $\tt CCGCAGCATCCTTGCCCCCCTGGCATGGGATCTGGGGCATCCTGCTTCTATTTGTTGGGCAGC$ ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCAGCTGGTGATGCGGTACTGGGAGCC CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCCACCTGCGCCACCCAGTGTGTG $\tt TGGAGCTGCTGACAGTGCTGTGGGTGCCTACCCTGGGCACGGACCGTCTCCTTGCT$ TTCCTCCTTACCCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT $\tt CCGGGCCCAGCTACAAAGAAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAG{\underline{\textbf{T}}}$ **GA**GGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCTCCCACTGAATTCTAAATCCTTAAC ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGGTATGGAAGTCATCTGTCCC TTCACTGTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC CTGACCACTCCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA GGCCCCAACCTTGCCTCACCACTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGGCTC

FIGURE 135

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP LAWDLGLLLLFVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP RALRLFSHLRHPVCVELLTVLWVVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA AGAAATTGCCAAACCATGTCTTTTTTTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTCATCTGGAATT TCACGTGGTGCTCCCGACTACTCACCCCGAGTGTAAAGAACCTTCGGCTCGCGTGCTTCTG AGCTGCTGTGGATGGCCTCTCTGGACTGTCCTTCCGAGTAGGATGTCACTGAGATCC CTCAAATGGAGCCTCCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT TCCCCACTACAATGTGATAGAACGCGTGAACTGGATGTACTTCTATGAGTATGAGCCGATTT ACAGACAAGACTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATCCATTT CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCCAGGCCATTAGAGTTAC TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGC TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTT GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC ACATGCCATTAT**TAA**CTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTCACAGTTATTATTATTTAAAATTA CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAAT ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQD FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEK EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD LVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG FSSKEIITFWQVMLRNTTCHY

FIGURE 138

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN PNKVDDLSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

FIGURE 140

CATTTCTGAAACTAATCGTGTCAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT ${\tt CAAGCTTATAGTTGAAATATTTTTCAGGAATTAC} {\color{red} {\bf ATG}} {\tt AATGACAGTCTTCGAACCAATGTGT}$ TTGTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT CAGATTCCGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTTGGTACTACAGAAGAGGAAAT CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAAACTATGAATTAC TGGAAAAAGAAGTAGAAAAAAGAAAGTAGCCTTACAAGAAGCCAAATTAAAAGCAAAGGGA TTGAATCCGGATGGAACTCCAGCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT GGTTCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGACATGG TCATAAAAGGAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT ${\tt C}\underline{\textbf{TGA}} {\tt CTTTCTTTTGAGCCTGCATCAGTTCTTGGTTTTTGCCTATCTACAGTGTGATGT}$ ATGGACTCAATCAAAAACATTAAACGCAAACTGATTAGGATTTGATTTCTTGAAACCCTCTA AAAATGCCCTAGCAGTATCTAATTAAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTACAAGGAAATAAAATACAAAT CTTGTTTTTCTAAAAAAAAAAAAAAAAAAGT

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ
SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSGHGRHRR

FIGURE 142

FIGURE 143

GGTCCTCGGAGTACAGCTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCCGATACAAGCACCCG TCTGAGGAGGAGCTTCGGGCCCTGGCGGGAAGCCGAGGCCCAGAGGCAGGAAAGAGCGGTG GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG AGACCTGCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACAT GCTGGGACCAGCCAAGGAGACTAACATTGCTGTTCTGGTGCCTGCTCACGGTGACCTTCT CCATCAAGATGTTCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGGTGAGCGC TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCCATGCTGGTGCAAGTGGTGCG GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC CCAGACCCACCGGGACGCACTGACCATGTCGGAGGACAGACCCATGCTGCAGTTCCTCCTGC ACACCAGCTTCCTGTCTCCCCTGTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC TTCCTGCACCAGCCGCTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA CTCTGGGCGCCTCTGGTTGCTGGTGCTGCTGCTGCGGCTGGCGGTGACCCGGCCCC ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCCTCAACTGCACACTTCTGCTCAAGACGC TGGGAGGCTATTCCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCCGACCCATCCTCAGCC AGCGCTGCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG GGCCCTGGGTGGCCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT GGACGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA ${\tt GGCTCC} {\color{red}{\textbf{TAG}}} {\tt CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCCTGGGGCAGCGGGACA$ CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCTTGGACGCCGAGTTTCTGCCTCA GAACTGTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGTTCTGAACAAATAAAGGAGCATGCC GATTTTTAA

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV VREETLELGLEPGLASMTQNLEPLLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFLTFPGLR LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT VVSLQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAARI AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLAGS

FIGURE 145

FIGURE 146

GGTTCCTACATCCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTT AATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGGCATGGCAAGGTTTGCTTAAAGGAGC CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG TATTCAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG TCCAGACGCTGTGCGAGCCTCACAGCCACGCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAA GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG GCAGGGAAAATTCTGAAAACACCACTGCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACTACGCTGTGCGTCTCCTGCGG $\tt CAGCCCTGCCAGGTGCTGTGGCTGATGCGTGAACAGAGTTCCGCAGCAGGAACAATGGACAGGCCCC$ GGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAA TAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA CATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCCAGAAAG TGCGGCTCATCTGATTCAGGCCAGTGAAAGACGTGTTCACCTCGTCGTGTCCCGCCAGGTTCGGCAGCGGAGCC CTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGAGAGGAGCAACACT CCCAAGCCCCTCCATCCTACAATTACTTGTCATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT GAGGAGTCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTTTGTTGAATGTGGATGGGGTCGAACTGACA GAGGTCAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT GTGACTGGTCCCCATCCTGGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTA**TAG**AATCAATGATGGGTCAGAGGAAAACAGAAAAA AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAAATTA AAAAAGCCCTAAAACAACTAAAATGATTGATTTGTATACCCCACTGAATTCAAGCTGATTTAAATTTAAAATTT GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTAATTTACAGCTAAAATATTTTTAAAATGCA TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTCAGAAGTTAAA

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT

APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SESLSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVVSRQVRQRS
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

FIGURE 148

FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT IFDYKHGYIASRVLSRRACFILKMDHQNIPPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

FIGURE 150

ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGCCAGCAGCTTCTCCAAGGCACG GGAGGAAGAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA GCCCCATCCTTGCCGTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGAGGAGTTTGGG GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAG**AATGAACC GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTCGTATTTGGA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPPITYSL CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK PVSELRANFTLQDRGAGPRVEMICQASSGSPPITNSLIGKDGQVHLQQRPCHRQPANFSFLP SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

FIGURE 152

GGTCCTTA**ATG**GCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG CTGTCCGGCTGGTCCCGGCCTGGCCGACCCTCACTCTCTTTGCTATGACATCACCGT CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAGA CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACCTGTCAGTCCCCTGGGGAAGAA CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCTGC AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT TGGAGCCAGAAAGATGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACACCC CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCCAACCCAACTCAGGGCCAC AGCCACCACCTCATCCTTTGCTGCCTCCTCATCATCCTCCCTGCTTCATCCTCCCTGGCA TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAAGGCTCCTGTGAGCACG GTCTTGATCAAACTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT GGCCTCCAGCAGATCATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG AATTTTTAAATTATTTAATAAGAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL HYDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHELWLLIKREFGFYSKSQYRTWQKKLA EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATG**AAG CTCTTGTGGCAGGTAACTGTGCACCACCACCTGGAATGCCATCCTGCTCCCGTTCGTCTA CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCCGCCCTCAGCCGGGCCCC AGAACTGCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCCTGCAGT TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC ACCCTGGAGCTGTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC CAAGCTGCGGGAGCTCTGGCTTCGCAACACCCCATCGAAAGCATCCCCTCTTACGCCTTCA ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA AGACATGCCCAATCTCACCCCCCTGGTGGGGGCTGGAGGGGGGGAGTGTCAGGGAACCACT TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCATG CAACTTGGCCCACAATAACCTCTCTTCTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC TGGTGGAGTTGCATCTACACCACAACCCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCATGCTCCCAT GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT CGGACTCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC $\tt CTCCCGCCACCCAAGGATCTCTGTCCTCAACGACGGCACCTTGAACTTTTCCCACGTGCTGC$ TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACTACAGCTTCTTCACCACAGT AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCCTA CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC AGCAGCAACAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTC ATGACCATATTAACTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC CTGGGGAACTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTCAGACCCA TACCAAGGACAAGGTACAGGAAACTCAAATA**TGA**CTCCCCTCCCCCAAAAAACTTATAAAAT GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAAATTTAAAGA CAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEMSGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHNPWNCDCDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

 $\tt CGCTCGGGCACCAGCCGCGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCC$ TTCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATG TGTCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGGAAAGAGGGGAAGTCGTGGGTTATACCAT $\tt TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT$ GTTGGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTCATGCTAAACCTGGGTTTGTCATCCAACTAA GATTTGTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTCGTGATGGAGACAAC CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGCCAGCTCCTATCCAGAGCATAGGATCCTCACT GCTCCTCATCCCCTTGTTTCCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG TGGGTACCAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTT TCTTTTGTAACAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGG AAACAGCCCATCTGCATAAAAGCCTGCCGAGAACCAAAGATTTCAGACCTGGTGAGAAGGAGAGTTCTTCCGAT CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAG TATGAGTGCATCTCACCCTTCTACCGCCGCCTGGGCAGCAGCAGGACATGTCTGAGGACTGGGAAGTGGAG TGGGCGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTTGC GCTGGCCGTGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG TTCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGG GAAGGTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATG AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCCTGCTTGATGCT GACATCGCCATCCTGAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAG GCCCTGGCTTCAAGAACGACACTGCGCTCTGGGGTGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAG CATGAGGACCATGGCATCCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCCTTC TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCCTTCCCGGGACGACCATCTCCTGAGCCACGCT GGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTCACCAAG GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAA<u>TGA</u>ACCATGCTCATGCACTCCTTGAGAAGTGTTTC TGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCCTGAAGTGTGATTTGGCCTGTGAACTTGG CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT ATATACAAAACCTCTCCACTCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTG ACCAGGGAAGATCTGGGCTTCATGAGGCCCCTTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSG
VVSVVDSLLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA AGCTTTCTTGCCTGCAGTGAAGCAGAGATAGATATTATTCACGTAATAAAAAAC**ATG**GGC TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT GACAGTGGTTGGGTGGGCCACCAGTAACTTCGTTGGTGCCCATTCAAGAGATTCCTAAAG CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAAGGGAAAAACTCTGACT AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAG AGGCCAGAGCAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC CTCGTTCCCCACCGGAACAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT CCTGCAGAGGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT TTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGCAGTTTTTCAAGGTGAATGGATTCTCT AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG AATGAAAATTTCCCGGCCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCACGAGTCTGG AGAACAGATGGGTTGAGTAGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA TATCAACATCACAGTGGATTTCTGGTTTTGGTGCA**TGA**CCCTGGATCTTTTGGTGATGTTTTGG AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTTCTTAGCAGAGCT CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG AGGGTTAAATATGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCT GTGAAGTGGTGTCAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA GGACACAGTGAACTTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTCGGCTGCAAAGGCAG CAGTAGCTGAGCTGCTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT TCCAGTGATGCCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGA TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 161

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKÇEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

FIGURE 162

CGTGGGCCGGGGTCGCGCAGCGGGCTGTGGGCCGCCGGAGGAGCGACCGCCGCAGTTCTC GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCAATG CCCGGCAGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG CCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA GATGGAGAAGGGTCTCAGCTCCACCATĆCGTGTGGTCGGCCACGTGCCCGGGGAATTCCCGG TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG GTCCTCCCCATCACAGAGTTCCTCGTGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC GCAACTTCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT CAAAGTGGTGGCGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTTGGGG CCCACCCTAATTCAGACCTTCCAAAAGATGACCGTGACCTTGAACTTCCTGGGGAGCCCTCC TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC CTGTGTCCGTGGCCACACCGTACAACCTGACCCACACCTTCAGGGACCCTGGGGACTAC TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG AACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT GGAGACTCCATCTGAGTACCTGGAAATTGTTCGTGAGAACCACGGGCTGCTCCCGCCCCTCT ${\tt ATAAGTCTGTCAAAACTTACACCGTG} {\color{red}{\bf TGA}} {\tt GCACTCCCCCTCCCCACCCCATCTCAGTGTTAA}$ CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCATT TGCGTGGGGCTGTTGGCCTGGATCATCCATCCTGTACAGTTCAGCCACTGCCACAAGCC CCTCCCTCTCTGTCACCCCTGACCCCAGCCATTCACCCATCTGTACAGTCCAGCCACTGACA TAAGCCCCACTCGGTTACCACCCCCTTGACCCCTACCTTTGAAGAGGGCTTCGTGCAGGACT TTGATGCTTGGGGTGTTCCGTGTTGACTCCTAGGTGGGCCTGCCCACTGCCCATTCCT CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCCTCCCCCAGACAGCCCTAC CTGTGCCAGAGAGCTAGAAAGATCATAAAGGTTAAAAAGGTTGTAC CACACACACAGAAATATAAACACATGCGTCACATGGGCATTTCAGATGATCAGCTCTGTA TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA AGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGCCCTGGATGGGGGGCCAGGACT AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG AAACTTTCACTGAGGAAAAGGCCTTGCAGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA GATCGAGACCACCCTGGCTAACACGGTGAAACCCCCGTCTCTACTAAAAAAATACAAAAAGTT AGCCGGCGTGGTGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG GTGCGAACCCGGGAGGCGGAGCTTGCAGTGAGCCCAGATGGCGCCACTGCACTCCAGCCTGA GTGACAGAGCGAGACTCTGTCTCCA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 164

FIGURE 165

 ${\tt MALSSQIWAACLILLILLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH} \\ {\tt FPICIFCCGCCHRSKCGMCCKT}$

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC CTGGATCTTCCACC**ATG**TTCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC AGTCTCCTTTGGTATCCGCAAACTCTACATGAAAAGTCTGTTAAAAATCTTTTGCGTGGGCTA CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC GGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTCGTCGAAGTGG TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC GGTCCTGTGGGGGTTAGGAGTGCTGATTCGGTACTGCTTCTGCTGCCGCTCAGGATAGCAC TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTTGCCAAAT AGCGCTGACAGCCATCATCACCATGACAGGGAAAACAGACCAAGAAATGGTGGCCATCT GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTCAGAGAGCCATGGTGAAGGCCTG CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA TGGTGACGTACCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC CATTGCCAGGCAGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC AAGGACAGGAGCCGCTCC**TGA**GCCTGCCTCCAGCTGGGGGCCACCGTGCGGGGTGCCAA CGGGCTCAGAGCTGCAGTTGCCGCCGCCCCCCCCCCTGTGTCCTTTCCAGACTCCAGGG CTCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACTCCCCA CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCCGCCACCCG CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAAATG GGCCGCTGACTGGGCCATGGGGAGAACGTGTTTCGTACTCCAGGCTAACCCTGAACTCCCC ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG AGGACACATCACGTTCAGTGTTTCAAGTACAGGCCCACAAAACGGGGCACGGCAGGCCTGAG TGA

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSGSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLLPLRIALAFTG
ISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRPRNGGICVANH
TSPIDVIILASDGYYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLPPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQQKLYSKMIVGNHKDRSRS

FIGURE 168

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA CGTCTCCTCCAGGG**ATG**GCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCTGGGCCTGGCTCCAGACACCTTTGA CGATACCTATGTGGGTTGTGCAGAGGAGGATGGAGGAAGGCAGCCCCCTGCTAAAGGAGG AAATGGCCCACCATGCCTGCGGGAATCCTGGGAGGCAGCCCAGGAGACCTGGGAGGAC AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG GCCCTGCAGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTTTCCG AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCCACAGATTTGGGGAGAAGAGGCGGGGCTGT GAAGACTCTGCTCTTGGCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCC**TGA**AAGTCCA ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL YMRHFPFKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

FIGURE 170

 $\tt GTGGCTTCATTTCAGTGGCTGACTTCCAGAGAGCAAT \textbf{ATG} GCTGGTTCCCCAACATGCCTCA$ CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG GTCGGTTCCGTTGGTGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC TATTGTCTGGACCTTCAACACCCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCA TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG CTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGGATATACAGCTCATCACT CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG GAACATGGGGAAGAGGTGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC CCATAATGGGTCCATCCCCCATCTCCTGGAGATGGGGGAGAAGTGATATGACCTTCATCT GCGTTGCCAGGAACCCTGTCAGCAGAAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT GAAGGTGCTGATGACCCAGATTCCTCCATGGTCCTCTGTGTCTCCTGTTGGTGCCCCT AGTACATTGAAGAGAGAGAGAGTGGACATTTGTCGGGAAACTCCTAACATATGCCCCCAT TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC TGCTCACGATGCCAGACACAAGGCTATTTGCCTATGAGAATGTTATC**TAG**ACAGCAGTG CACTCCCCTAAGTCTCTGCTCA

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

FIGURE 172

 $\tt CTGGTTCCCCAACATGCCTCACCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC$ TCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA GATGGAGGCTACTCCCTGAAGCTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGT GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT GGGGCAAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAAACTTCTCAAGCCCC ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCTCCATGGTCCTCCT GTGTCTCCTGTTGGTGCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC TGAAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGTGGACATTTGTCGGGAA ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACCCAAGGCTATTTGCCTATGAG

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATCGAAGATGCTGCTC
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTTCCTGTTCAATAAATGATTACCCTTTGCACTTAA

FIGURE 174

MKMLLLLCLGLTLVCVHAEEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ IHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

FIGURE 175

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATAGTGAA
AAAATGTGGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV KHCTDQISFKKRLSLKKSWWK

FIGURE 179

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK YKSSQKQHSPVPEKAIPLITPGSATTC

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCC AGTGTTTCCCACAGCCCCCAAAACGGAACTGGTTTTTGGGGTCACCTGGGCCTGATCACTCCT ACAGAGGAGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT ATGGCTGGGTCCCATCATCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCTGAAGCCCTGG CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC GCCCGCCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG TCAGGAGAGGCCCAGTGAATATTTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGCGGCGC TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG CACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG ATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC CTGGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC TTCCTGACCATGTGCGTGAAGGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG ATGCTGCACCCAGACATTGTTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTC CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG CGTTGATGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAA TTGATCATGCGCGCCGAGGGCGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA GTGACTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEGILLSGGDKWSRHRRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFFKDKAKSKTLDFIDVLLLSKDEDG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

FIGURE 183

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDSREISFQLSAPHEDARLTPEELERASLLQILPEMLGA ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

FIGURE 185

FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

FIGURE 187

CGGCCACAGCTGGCATGCTCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT**ATG**AACACGTGGCTGCTGT TCCTCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT GAGTATGTCCCCCACCCTAAGCCCCCGATCCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC TTACACCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT GTCCCCGTGTGCGCATGTGTTCAGGGTCAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAAC GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT AACAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTTGGTCATGGTGGGCCAGCTGCAAAGCG TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC ACGGCCAGCAGTGTGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTCAGGCTCTGGGCTCACCTCCATCTCCAGA GCATCCCTGCCTGCAGTTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCCACCAAGA GCCTCCTTGTTCATAACCACAGGTTACCCTACAAACCACTGTCCCCACACACCCTGGGGAT GTTTTAAAACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA ATTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDFLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ AQQEAELTPRPAGVVPGA

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAG ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCCACACAGAAGG TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG GATGGAGATACCAACACCCCCCCCCCGGGGGGTGGTACAATACAACTGGGAGACTGGGGATGA $\verb|CCGGTTCTCCTGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAG|\\$ AACCAGGGGAGAGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACTGAGCGCCTTTGCTG CTGTTTCCTCTGTCAGGTCTCCTGGGGATGGTGGCCCACATGATGTATTCACAAGTC TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAAC CCGAACTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG TCTCTTGAGCATGGTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC CTAAGGGATTCCTGGGTGCCACTGCTCTTTTCCTCTACAGCTCCATCTTGTTTCACCCAC CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC TAAACCATGGAGATAAAAAGAAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP VSLDGDTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLETTCLE LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

FIGURE 191

AACTGGAAGGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG TCTTTATGTCTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCCAATGGCATTACAGAGAAAG CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCGCCGTAGATTCAG GACATTCGCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT $\tt CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT$ ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA TGTTGAAGGGCACAAGAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

FIGURE 192

 ${\tt MWLPLGLLSLCLSPLPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQGSGME} \\ {\tt HRNHLCFCDLYDRATSPPLKCSLL}$

FIGURE 193

CCGCC ATG GCAGCCATCAAAGCTTTGATTAGTTTGTCCTTTGGAGGAGCAATCGGACTGATGTTTTTTGATGCTTGGATGTGCCCTTCCAATATACAACAAATACTGGCCCCTCTTTGTTCTATTTTTTACATCCTTTCACCTATTCC ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTC TTACAACGGGCATTGTCGTGTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA GCTTGTGCACTTGTTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTTGGAAG ${\tt CAATGACGACTTCAGCTGGCAGCAGTGG} \underline{\textbf{TGA}} {\tt AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTT}$ GTTGGCCATTCACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA ${\tt GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT}$ CTTTTGGAAAAGCTTGACTGATTTCACACTTATCTATAGTATGCTTTTTTGTGGTGTCCTGCTGAATTTAAATAT TTATGTGTTTTTCCTGTTAGGTTGATTTTTTTTGGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA TTTGCATTGGTTAGGAATTCAGAATTCCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAAATTATT CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG AAAATTATTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA GGGAAGAATGACATTGAAATTCCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC AACCTTTCATGTTTTACCCTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACTTGTTAGTCTTACAGATAATTCATGCA TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAG TATTTTTAAGACAAGTTTCCTGTATACCTCTGAACTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT ACCTGACCAAAAAATTCCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA TTTATGAAGTTTATTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTCAGCTTTTTTACTAAAGATGCCTAA AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG ${\tt GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTCAGATTTCAAGAGGAA}$ GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACTCTTTGTGCTTGTGATCTACTGGACTTT TTTTTTGCAGGAAGTGCATTCTCTGGTCCTTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCACTGCTACTG TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT GTGTCTTTGACCTTGTATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT AGATTTTAAATATCTATTTTAAAAAAAAA

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM SNACKELAIFLTTGIVVSAFGLPIVFARAHLIEWGACALVLTGNTVIFATILGFFLVFGSND DFSWQQW

FIGURE 195

GCTGATTTGGGTCTGCCATTGACAGAATGTCAAATAAAAAGGAATTAGCTAGAATATGACCATTAAATGTGCTT CTGAAATATATTTTGAGATAGGTTTAGAATGTCA

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSVDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLTSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDAVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVIINEYGSCTCHQQPARECEV

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATCGGCCGTCCTGGGCCGGGTCCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCGCGCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCCGCAGCATCTCGGCTCTGGCCGACGTTCACGCGCGACGAGGACCTGG
CTGTTTTCCTGGCGTCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGCCTCTGAGCGTG
GGCCCCGAGGACTGCGCGGACCCTTCGGCTCTGCGCCAACGCGGAGGCCCAGCCGTG
GATCTGCGCGGGCCCCTCCAGCCCCT

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR AGRLRFHGPGALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 199

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP FARDAVKKCFAVCLA

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCTCCAGTCACCCTCCCGCCGTTACCCGCGGCGCGC $\verb|CCGAGGGAGTCTCCTCCAGACCTTCCCTTGCTCCAAACTAATACGGACTGAACGGATCGCTGCGAGGGT|\\$ AACTGATCAAGTACTTTGAAAATCACTTCGAAATTTATCTTGGTGTCCTTCATACTTGCTGCACTGAGTCTTTC AACCACCTTTTCTCTCCAACTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT TATATAAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT TTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTGCAGAGAATCATGGGATTGTTGC AAATGATATGTTTGATCCTATTCGGAACAAATCTTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT GGGAAGAAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA ACAGATGTAAAAATACATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG AGTTGCCAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG ATGACATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTCAGATATTGACAAGAAGTTA GGATATCTCATACAAATGCTGAAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATACCCTGATTG ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT CATCCTAATCTTACTGTTTACAAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCA ${\tt ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC}$ TCAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA CAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA CTATACTCCTCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCATACCCTTATTTCATAGGGGTC TCTCTTGGCAGCATTATAGTGATTGTATTTTTTGTAATTTTCATTAAGCATTTAATTCACAGTCAAATACCTGC $\tt CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCC\textbf{TAA} TGTTACTTTGAAGTGGATTTGCATA$ $\verb|TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCTGGGAAACCAGTTCCAAACATCTGC|$ ${\tt ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATTTTAGCAACTTTGCACTATGTAAAGTACCTTATAT}$ ATTGCACTTTAAATTTCTCTCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC TTGATTGAAAATGACAACTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACTGAAGGAAATT TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGA AAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT ATTCGTTCTAAATATTGTTCTGTCATAAAATTATTGTGATTTCCTGATGAGTCATATTACTGTGATTTTCA TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA AAATCAAATTGGATAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTCGCTCCCACGGGCGGACCTTTGTAACTGCGGGAGGCCCAG AGAGAGGCCAAGCCCTTGCCTTGGGTCACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA CCAGATCCAGAGGCAACAGGGACATGGCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA AGCAGCCACCACCACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCTGACGTTGCC CCTGCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCCATGTTGAGGAAACTGTT TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG TGGTCTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCTGGAC TGA TGAGTTTGCTGTATCAACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAAGAAGAAGAAG CAGGCTGGCATGTTCACTGGGCTGGTGTTACGACAGAGAACCTGACAGTCACTGGCCAGTTA TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEQPPPTPV SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELILDL KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLSSFTTSLRSWMPVVVVVSFILDI VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLLRLKQMNVQLAAKIQHLEFS CSEKPLD

FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGCAGTGGAGCACCCAGCAGCCCCAAC**AT** $\underline{\textbf{G}}$ CTCTGTCTGTGCCTGTACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTT $\overline{ ext{TG}}$ ĀGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC CAGGAATTCTCCACCTACCGCCAGTGGAAGCAGAAAATTGTACAAGCTGGAGATAAGGACCT TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA CATGGATAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCCACCTGGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG GAGACACCTGGTGCCGGATGAGTTCACAGTGGAGAGGCAGACGGGGATGTGGTGGAGACACCTGCACGGCCCCCCTGGAGACCCTGGTGCAGGGCAGGGGCCCCCCTGGACAGGCTCAAGGTGCTCATGCAGGCCATCCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAACCCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCCTTGTTGGTAGTGACCAGGAACACCTCTGGCAGGGTCCTTGGCA GGGGCCATCGCCCAGAGCAGCATCTACCCAATGGAGGTCCTGAAGACCCGGATGGCGCTGCG TGGTGCTCTGAGCTGGCCTGGACCCTGTCAGGATGGGCCCCACCTCAGAACCAAACTCACTG TCCCCACTGTGGCATGAGGGCAGTGGAGCACCATGTTTGAGGGCGAAGGGCAGAGCGTTTGT GTGTTCTGGGGAGGAAAAGGTGTTGGAGGCCTTAATTATGGACTGTTGGGAAAAGGG CTGGGGGCGTGGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTCCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW WRHLVAGGGAGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMAL RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 207

GGAAGGCAGCCAGCTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCAT CAATTGCACTCATCATTGGCTTTGGTATTTCAGGGAGACACTCCATCACAGTCACTACTGTC GCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAA ACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCA AAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTTT GCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGC TGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATA AAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTG CGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCA GGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCA TGAAGGTTGTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAA AATGACATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAAGGCGGAG ${\tt GGGCACTTCTGCCTCTAGCCCTTACCTGATGCTAAAA} {\color{red}{\textbf{TAA}}} {\tt TGTGCCTTGGCCACAAAAAAG}$ CATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTCACCACCAGATATGACCTAG TTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCA GACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAG GGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTG TAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCA CAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGG GGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCT TGGCTTCTCTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAA ACAGAGCAGTCGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAA

FIGURE 208

MASLGQILFWSIISIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD QGANFSEVSNTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRR SHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCTATCATCTTCA CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG GTTTGGGAT**TGA**AGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCCTCCCAGGAAGCCCCT GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT CCTCCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG AGTTTTGTAACCTTCAAGTGCTGTTCAGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA AAATCCTCACTGCCAGCCCCTCTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCACCTGAC

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLLQEMKTLFLNTEYLMPFL LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS RHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTCGAAAAGATTCCGCAATAAAACT TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGCATCCTCG TTGCTGGTATCACTGCAGTGCTTGTTGCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT GCTGAAGAACACTTTCATTTTGTAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG AATCTAATGGAACTTCCTGTCGTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCA TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG GGGACTGCTGCCC**TGA**GGTCCTGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTG AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTTCCCTGCTCTGCCCCGTTTAACTGC CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK TLGGVIFRKFECANVNSLTPTSAPTTSHNVGSKASLYLLALASLLLRGLLP

FIGURE 213

GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCCATGGTCCCCGCCGCCG ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA ATGATGCCATGGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA GGGGGTTGTGATTAATGCCGGAAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCCTCACCCACAG CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCT TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGCACCTATCAACAATGTC CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAGTCTCTGTACTGACACCAACTGT GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG AAGCAGTCCCAGCCTGCCACCCGCCAGCCCTGCCCAGCCCTGGCTTTTTGGAAACGGGTCA GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTCACAGAGATGCAACCAATA GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACTGGAAAA CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP SPSPTAMPSPEDLRLVLMPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC TYQQCPCNRLREECPLDTSLCTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

FIGURE 215

GCGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAGTGCGAAGAGCA GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTCACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCT CAGTTCCATGGCAAGTGGCCCTTCTCCCGGTTCCTGTTCTTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCT CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCCATGTACCACA CCTGTGTGGCCTTGGCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC CCTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTG GTGTGGTGGCTGGCCTGTGGAACCAGCGGCGCTGCCTCACGTGCGCAAGTGCGTGGTGGTGGTCTT GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCCTGGATGCCCATGCCA TCTGGCACATCAGCACCATCCCTGTCCACGTCCTCTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG ${\tt AAGGAATCAGAGGACAAGTTCAAGCTGGAC} {\tt TGA} {\tt AGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCCC}$ ${\tt GCCCTGCTGGCCTCCCTTCTCCCCTCAACCC} \overline{\tt CTTGAGATGATTTTCTCTTTTCAACTTCTTGAACTTGGACATGA}$ ${\tt AGGATGTGGGCCAGAATCATGTGGCCAGCCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG}$ AAGGCCTCCCAGCATCTGGGACTCGAGAGTGGGCAGCCCCTCTACCTCCTGGAGCTGAACTGGGGTGGAACTGA GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCCTCCCACCAGCCTCCCCCACATCCCCAGCTG ${\tt CCTGGCTGGGTCCTGAAGCCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC}$ GGTTCACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTCACCCTGACC GTTGCCCTAGCCAGGTTCCCAGGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCCAGGGCA AGGATCCTGTGCTGCTGTTGAGAGCCTGCCACCGTGTGTCGGGAGTGTGGGCCAGGCTGAGTGCATAGG TGACAGGGCCGTGAGCATGGGCCTGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGT TGCGCGTGCTGGTGGGCATGTGAGATGAGTGACTGCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGAT GAGGGAATCCTGTCACCATCAATAATCACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGC CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGCCCCTTTTGCCCGCCTCCTGCAAAC TATGGCTGGGTCTGGTTTCTTCCCTTCCCAGAGGGTCTTACTGTTCCAGGGTGGCCCCAGGGCAGGGGGCC ACACTATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCCTGCCCCACAGG AATAGAATGGAGGGAGCTCCAGAAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTTG CTCTGCCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGGGGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG GCCTGCGCTAGCTTCTTTTGATACTGAAAACTTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA TTCCAAGCCTCAAAAAAAAAAAAAAAAA

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLLQGLSLLELLDFPPLFWVLDAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGA $\mathtt{CT} \underline{\mathbf{ATG}} \mathtt{GCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTG}$ CTGATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGAC TGGATGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGG CAGGCTGTCCCTGCAGGATCCCACGTACGGCTGAATCTTCAGACTGGGGAAAGAGAGGCAAA ACTCCAATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCA ACACCTACACATCTCAGGATCTCAAGAGTGCACTGGCAAAATTCAAGGAGGGGGCAGAGATG GAGAGTTCAAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGA GGAACTGAAGAAGACTTTGATGAGCTGAATGTTGTCATTGAGACTGACATGCAGATCATGG TACGGCTGATCAACAAGTTCAATAGTTCCAGCTCCAGTTTGGAAGAAGATTGCTGCGCTC TTTGATCTTGAATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCTTTGGTGG TCTTCAAGTGGTGATCAATGGGCTGAACAGCACAGAGCCCCTCGTGAAGGAGTATGCTGCGT TTGTGCTGGGCGCTTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGA GCCCTGCAGAAGCTGCTGGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAGGT CCTGTTTGCACTGTGCTCCCTGCTGCGCCACTTCCCCTATGCCCAGCGGCAGTTCCTGAAGC TCGGGGGGCTGCAGGTCCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTG CGCGTGGTCACACTGCTCTACGACCTGGTCACGGAGAAGATGTTCGCCGAGGAGGAGGCTGA GCTGACCCAGGAGATGTCCCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAG GCCTGTGGGAACAGGGCTGGTGCGAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGAT GCCCGTGAGAAGGTGCTGCAGACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCG TCAGGACCCCAGCTCGGCAGGACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCA GCCTGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAAC AGCTTGCTGAAGGAGCTGAGATGAGGCCCCACACCAGGACTGGACTGGATGCCGCTAGTGA GGCTGAGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCT

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL
DAEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRTLVQEKGTEVLAVRVVTLLYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

Important features:

Signal peptide:

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCCTTCCC CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCTCGTGGGGTCGCGTTGCCACCCCA CGCGGACTCCCAGCTGGCGCGCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCC $\tt TTCCCACCTGACCAGCC \underline{\textbf{ATG}} \tt GGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGC$ CCGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT CGCAGGGGCATTTTTCTGGCTGGTCTCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG TCCATGTGACCGACCGGTCAGATGCCCGGCTCCAGTACGCCTCCTGATTTTTGGTGCTGCT GTCTCTGTCCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT ATGTTTCTGGTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCCTATTACTTCCTGACTTC AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTTTTTTGATG CCTGTGAGAGGGGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG GGACTGACATTCCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC TCTTGTGTAAGGAC**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG CCCATGACTGAGCCCAGCCCCAGCCCGGGTCCATTGCCCACATTCTCTGTCTCCTTCTCGTC GGTCTACCCCACTACCTCCAGGGTTTTGCTTTTGTCCTTTTGTGACCGTTAGTCTCTAAGCTT TACCAGGAGCAGCCTGGGTTCAGCCAGTCAGTGACTGGTGGGTTTGAATCTGCACTTATCCC CACCACCTGGGGACCCCCTTGTTGTCCAGGACTCCCCCTGTGTCAGTGCTCTCAC CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC CTGGGGTGCCCTCTGATGTCCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC GAGGTGGGGGGGGGGGGGGTATATTGGAACTCTTCTAACCTCCTTGGGCTATATTTTCTC TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT GGGGGAAAGGAAGTGCATGTTTGGGAACTGCCATTACTGGAACTAATGGTTTTAACCT CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG ATCATTTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATCAAGGTGATTGTGATTTTGACT

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIILVAGAFFWLVSLLLASVVWFILVHVTDR SDARLQYGLLIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS FGIISGVFSVINILADALGPGVVGIHGDSPYYFLTSAFLTAAIILLHTFWGVVFFDACERRR YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

FIGURE 221

FIGURE 222

GACCGACCGTTCAGATGCCCGGTTCCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCCCCCCTTTCCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCATCTTTGGTCCATGTGAC
CGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGCTTCCCCCTTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCCC
CAGNTGGCGCGCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCTGCTCCTGGCCTCTGTGGTCATCTTTGGTCCATGTGA
CCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGAGGCAG TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCCACCCC TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT $\tt CCTTGTGGTTCCTCTACCTGGGGAAATAAGGTGCAGCGGCC{\color{red} ATG} GCTACAGCAAGACCCC$ CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGGTCACAGAGCATGTT CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA GGACCTGGGAGCTGGGGCCGGGAAGACGCCCGGTCGGATGACAGCAGCAGCCGCATCATCA ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC CAGCTCTACTGCGGGGCGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG GAAGAAGTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCTGGCTACTCCCACCCTGGCCACTCT AACGACCTCATGCTCATCAAACTGAACAGAAGAATTCGTCCCACTAAAGATGTCAGACCCAT CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGGCCTGTGGTCTGCAATGGCTCCC TGCAGGGACTCGTGTCCTGGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTAC ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCAT CCCAGGACTCAGCACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG ACCCTCATTCCTTCCCAGAGATGTTGAGAATGTTCATCTCTCCAGCCCCTGACCCCATGTCT CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG GAACAATTTCCAAAACTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA CTGAGAAGTGGAAAAAAAA

FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCPSAGTKCL VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGRDSCQGDSGGP VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCCGCCGCCCAACTTGAGGACCGGCCGCGA CAAGCCGCAGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC TGCTGGCTGTAGCTGTCACCGGTGCCGTGCTCTTCCTGAACCACGCCCACGCGCCGGGCACG GCGCCCCACCTGTCAGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGGA AAGGGCGGACAGCTCGCACCTCAGCATCCTCATTGACCCGCGCTGCCCCGACCTCACCGACA GCTTCGCACGCCTGGAGAGCGCCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC ACGGCTCCGTGAACTTCTTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCACC TGCTGATGGCGTCCCCACCCCCCCACTTTGGCAGCTCCCACTGTGACTCCTGCTGCTGCTGCTGCCCACTGTGACTCCCGGTGCCAAACTTGACGACCCCAACTCCTGCCCCACCCCCACTGTGACTCCCGGCTGTGTTTTGCCGTCCCCTGGCCAGGATGGTGGAGTCTGCCCCAGGCACCCCTTGCCCCCGGCCAAATACCCCGGCATTATGGGGACAGAGAGCAGGGGGGCAGACAGCACCCCTGGAGTCCTC GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCCTAGGCAGGGCTGGGATGAG AACAATAAATTTGACTTGGCACCACTGGGGGTTGGTGGGAGAGGCCGTGTGACCTGGCTCTC TGTCCCAGTGCCACCAGGTCATCCACATGCGCAG

FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDTLADQLPRLLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRDSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASYADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCCAGGGCAATCCGACCACATTTCACTCT CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCACGAGGGACATGCTGGATGATG ATGGGGACACCACGATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC TTTGTGCTTGGTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACC AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAGATTAGGAAATACGTCC CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC TGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGAACCCCCTTTCACTTCTGAACTG TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA TGGTGAAGCCAGAGACCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGAC**TGA**TTCGCC CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG TTTCCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTCAT GTCTTCCTTACACTTGGTGGAATAAGAAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC ATCCTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC AGCAAATACACAAGGAATTCTTTTTTTTTTTTTCAGTTCATACTAGTCCCTTCCCAATCCAT CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG AACTACCGTCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD CKELKRCVCERRAGMVKPESLHVPPETLGEGD

FIGURE 231

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGCTC CCCGAGCCCTCCGGATCCGCCCCCTCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT $\mathtt{CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCC}$ ${ t TCGGTGCTGCGGCCCGCAGGGCCCGTGGCCGTGGGCATCTC}{ t CCCT}{ t GGGCTTCACCCTGAGCCT}$ AGCTGCCGCCGCGCGAACACCAACGCGGCGCGCCCGAACTCGGTGCAGCCCGGAGCG GAGCGCGAGAAGCCCGGGGGCCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTACATCAGCA ACGCTGGGCGTGGACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC GGGCGCACGGGGCCCCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC CCATTGGACACCTGCCGCGCGCCCCCCTGCTGGAGCACCACGGCGACGACTTTGAC TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG GAGAGCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGGTGCTGCTGTCGCGCATGCTG CTGCAACAACTGCGCCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA CGAGTGGCTGGGTCGCTGCATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT TTCCGAAGTGCCCTGACAGCCCACCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA AGCTTTCGCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT CCAGCACCATCCCGCCCGGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCACTGCGTGGGGCTGACCGGGCTG ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG CGGCTCCAGAAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCGGGGTATGGA ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCCAGGGAGGCCGCCGGCCCCTCACTC GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCTGTGCCCTATGTCACT GAGGCCTCACGTCTCACTGTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCTGG CTTCTTGGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC TGCTGCTACTGTATGAGCCGCGCCAGGCCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG TGTGCAGACAGCCGCACCCTCACCACTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC TGGACACACTGTTCCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA CCCAGGTGTGGCCCCACACAAGGGCCTGGGCCCCAGAGCTGGGCCGTGACACTGGCCGCT TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGGGCCTGCTGGAGAGCCTGGATGTGTACGAGCT ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG ${\tt CAACAGCACC} \underline{{\tt TGA}} {\tt CCCCACCCTGTCCCCGTGGGCCGTGGCCATGGCCACCCCACCCCACTT}$ AAGCTGGCCCACTGGTCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG GGACGTGCCCCAGAGCCACCTCTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT GCTGATTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA **AGTGTGGAAAAA**

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPPRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLLE
QHGDDFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCILDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALTAHPVRDPVHMYQLHKAFARAELERTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGPGPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLESLDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTOLAMLLFEOEOGNST

FIGURE 234

GCTCTGGCCGGCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT TGGCAAGCGCTGGCCACCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT AGCTATTAGCCAATTCGGCAGGGCCCGCTTTTTAGAAGCTTGATTTCCTTTGAAGATGAAAG ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG GCGGAGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT TTCCCCGCCCTGAGACCCTGCAGCACCATCTGTCATGGCGGCTGGGCTGTTTGGTTTGAGC GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCCGCGTCCGCTGGGA ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACTTGTATGAGAAGAACCCA GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC $\tt CTTCCCATCATGGAATCCAACTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG{\bf TG}$ CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER LVKYREANGLPIMESNCFDPSKIQLPEDE

FIGURE 236

FIGURE 237

TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTCATCACCCCGCTGCCTTCCGGGG ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTCGGAGCTTCAGCGGGAAGGAGTG ${\tt TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA}$ GCTGCACCTGTCATTCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCCACCCTTCCTGC AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCCTCTCTCAACTTCATCGA CTCCACCAACACTCCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC CCCTGGAAGAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA TCGCTTGTTCCACACCAGCTACCACTCCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC ATCACGGGGCAGGGAAAGAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCTCACGGA GCCCTGCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACCTACAACCAGGACA ACGAGACATTAGAGGTGCACCCACCCCGACCACTACATATCAGGACGTCATCCTAGGCACT CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACTCTCGAAACCT CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCAGTGCCCTTCCTGC ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC AACACCCACCATACCGGGCCTTCCCGGTGCTGCTGGACACCGTACCCTGGTATCTGCG GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAAACCAAGTTACATCC ACTACCAGCCTGCCAGGACCGGCTGCAACCCCACCTCCTGGAGATGCTGATTCAGCTGCCG GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCCATCTGTCCTCAGCGCCCTTGTGCCCA GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCCACTC**TGA**TT ${\tt CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT}$ TCTGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTTGAATTTAA CTTAGAAATTCATTTCCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA AGTGGTCGGTGGCTGTATTGGACAGCACAGAAAAGATTTCCATCACCACAGAAAGGTC GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTGTAGTGGA

FIGURE 238

MPLALLVLLLLGPGGWCLAEPPRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSISWELRQTLSVVFDAFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTVPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

FIGURE 239

FIGURE 240

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

FIGURE 241

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC ${\tt TCTAGAACCCGACCCACCACCACCACCACCACCACCTGAGCCTGAGCCTGAGCCAAGG}$ CGTCCAGTGGTCCTTGCTTCTGGCTGTCCTGGTCTTCTTCTCTCTCGCCCTTGCCCTCTTTTA CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCCACACCA ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACTGTC ACCCAGAGGCCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCCACAATCATGGAAGAGCC AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG GTGTCAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAAGTCA GCACAGAATGCTGGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCCACCCCCTGCCCCTTTCCAG AGCCCCACGACGCAGAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCCAACCTCACTCTC TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG TGCCCCAGCAGCAGCTCCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT GCCGTGGTGGCCAACGGGGCATCCTGAACAACTCCCACATGGGCCAGGAGATAGACAGTCA CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCCTGGAAGGCAC CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT TCTGGTTCAGGCACAGACCCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCCATGAGCGCTTTTCTGAT CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC CCGGAACTGCCAAAGCCAAGAAC**TGA**CCGGGGCCAGGGCTGCCATGGTCTCCTTGCCTGCTC CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCCTGAGGCCAGG GATTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAACACTCTT CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCCTTG TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACTGATAATAATACAAATGATTGTT

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
NQSEWDRLEHFAPPFGFMELNYSLVQKVVTRFPPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 243

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT RDCTIPAYYKRCARLLTRLAVSPVCMEDK

FIGURE 245

GGGCTGGGCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG CCCGACCCGGCCGCGCCCAGCCCACCATGCCACCGCGGGGCTCCGCCGGGCCGCCG CTCACCGCAATCGCTCTGTTGGTGCTGGGGGGCTCCCCTGGTGCTGCCGGCGAGGACTGCCT GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT CCTCTTTGTTGCTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCCTGTTGCTACCTGT ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTTGAAGGCCAGGAGATTCCAATGACAGGCATC CCAGTGCAGCCAGTATACCCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC CAGTCTACAACCCTGCAGCTCCTCCTCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC TGAGGAACCAGCCATGTCTCTGCTGCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA ACTATGAGGGGTTGGGGGGGGCTTGGAATTATGGGCTATTTTTACTGGGGCCAAGGGAGG GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG TCCGTCAGCAGCTGGCAGTAGCCCTCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTFCCGTCYHRYC CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQLQSP FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGCCAAGGGTGAGGGCCGCCCAGAA CCCCAGGTAGGTAGAGCAAGAAGATGGTGTTTCTGCCCCTCAAATGGTCCCTTGCAACCATG TCATTTCTACTTTCCTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTCATGGTGTCAGAG CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAATAAAATACGACTTC CTGAGTACGTCATCCCAGTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC TTCTGGGGAACCACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCCTGCA TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG AAGAACCCCTGCAGGTCCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAG CCCCTCCTTGTCGGGCTCCCGTACACAGTTGTCATTCACTATGCTGGCAATCTTTCGGAGAC ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA AGATGAGCACCTATCTGGTGGCCTTCATCATTTCAGATTTTGAGTCTGTCAGCAAGATAACC AAGAGTGGAGTCAAGGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTCAGCATACCGTATCCCC TACCCAAACAAGATCTTGCTGCTATTCCCGACTTTCAGTCTGGTGCTATGGAAAACTGGGGA CTGACAACATATAGAGAATCTGCTCTGTTTGTTTGATGCAGAAAAGTCTTCTGCATCAAGTAA GCTTGGCATCACAGTGACTGTGGCCCATGAACTGGCCCACCAGTGGTTTGGGAACCTGGTCA CTATGGAATGGTGGAATGATCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTG TCTGTCAGTGTGACCCATCCTGAACTGAAAGTTTGGAGATTTTCTTTGGCAAATGTTTTGA CGCAATGGAGGTAGATGCTTTAAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTG CTCAGATCCGGGAGATGTTTGATGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATG CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAAGTGGTATTGTACAGTATCTCCAGAAGCA TAGCTATAAAAATACAAAAAACGAGGACCTGTGGGATAGTATGGCCAAGTATTTGCCCTACAG ATGGTGTAAAAGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT TCCCCTAATAACCATCACAGTGAGGGGGGAGGAATGTACACATGAAGCAAGAGCACTACATGA AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTCATCACC AGCAAATCCAACATGGTCCATCGATTTTTGCTAAAAAACAAAAACAGATGTGCTCATCCTCCC GAAACTCAATTCAAGGCCTTCCTCATCAGGCTGCTAAGGGACCTCATTGATAAGCAGACATG GAAAGCTTTAAGGGAGATAAAATAAAAACTCAGGAGTTTCCACAAATTCTTACACTCATTGG ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAA**AAA TTCCTCCCTTGCCCGGTTCCTGTTATCTCTAATCACCAACATTTTGTTGAGTGTATTTTCAA ACTAGAGATGGCTGTTTTGGCTCCAACTGGAGATACTTTTTTTCCCTTCAACTCATTTTTTGA CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTCATGAATGGGCTA TCGCTACCATGTGTTTTGTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTTGGGT

FIGURE 248

 $\verb|MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH|$ YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE HPPOEOIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTOFEPTA ARMAFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL WLNEGFAKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG FCSRSQHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL TGLLKGTHTAVSSNDRASLINNAFOLVSIGKLSIEKALDLSLYLKHETEIMPVFOGLNELIP MYKLMEKRDMNEVETQFKAFLIRLLRDLIDKQTWTDEGSVSEQMLRSELLLLACVHNYQPCV QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC RTQNKEKLQWLLDESFKGDKIKTQEFPQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS SSIAHMVMGTTNQFSTRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

CAGCCACAGACGGGTC**ATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG AGGCCAAGGACCAGGAGCCCCGCTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC TTGGGCCCCACAGCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGG AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGA ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT GTGTCAGGAGACGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCTCCTGGG GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG CAGCGTTCTGCTGAACTCCCTCCTCAAGCTGCCCCTGTCCCAGGAGACCGGCAGTGTC CTACCTGTGTGCAGCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG GGCGCCACTCATTGTTATGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAAT GAGCATTCAGGGCTGCCTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG GGGGCTGAGGCCTGGAGTCTCTCACTTGGGGGGTTGGGCACTGGCCCCAGCGCTGTG GTGGGGAGTGGTTTGCCCTTCCTGC**TAA**CTCTATTACCCCCACGATTCTTCACCGCTGCTGA CCACCCACACTCAACCTCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGGACGCTGGAGGAGTG GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC PSC

FIGURE 251

CAGG ATG AGGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG ${\tt CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG}$ CCTCAAAGGGGATGCGGGAGAGAGGGGAGACAAAGGCGCCCCGGACGCCTGGAAGAGTCG GCCCACGGGAGAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCGTCAT GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTCGCC GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA CGCCCAGCTGTCCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCGTGTCTTCATCGGCATCAAC GACCTGGAGAAGGAGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA GTGGCGCAGCGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG GAGAACATGTGAGCCTCAGGCTGGGGCTGCCCATTGGGGGCCCCACATGTCCCTGCAGGGTT GGCAGGGACAGACCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG AAAGTGTTCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC

FIGURE 252

MRGNLALVGVLISLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG PTGEKGDMGDKGQKGSVGRHGKIGPIGSKGEKGDSGDIGPPGPNGEPGLPCECSQLRKAIGE MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS GGWNDVACHTTMYFMCEFDKENM

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG CACCAGTGTGTGAGGGGAGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC CAGCCCCGGCACAGAGCTGCTCCACAGGCACCATGAGGATCATGCTGCTATTCACAGCCAT TTCCTGGCGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCC CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC TTTA**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT ATCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTTCCCTGTTCTTGT AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT CCTACATTAAAAATATAATGTCTCTCTATTCCTCAACAATAAAGGATTTTTGCATATGAA

FIGURE 254

MRIMLLFTAILAFSLAQSFGAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK ALSQASTDPKESTSPEKRDMHDFFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCGCTGTGCCTGTGCC CGCGCTGTCGCCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC CCATAGTTGCTGCAGGAGTGGAGCC**ATG**AGCTGCGTCCTGGGTGGTGTCATCCCCTTGGGGC TGCTGTTCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG CTGCTCAGCAAATACCAGCACAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG $\texttt{AAACCT} \underline{\textbf{TAG}} \texttt{ACTCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACTACAG}$ GCATGCACCATGGTGCCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA ACAACACGTGGGTTCCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC TCCACTGGGAACACAGCTCTCAGCCTTTCCCACCTGGAGGCAGAGTGGGGAGGGGCCCAGGG CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG TGGAAACTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAAATC CAAGTTAGCCGGGCATGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG GTGGAGGTGGGGGGGGGGGGGAGGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLEELLSKYQHNESHSRVRRAIPREDKEEILML HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

FIGURE 257

FIGURE 258

 ${\tt MGSGLPLVLLITLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTS} \\ {\tt VTLHHARSQHHVVCNT}$

FIGURE 259

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFSVENECLVDLCLLRICYKLSGVPNQCRVPLP SDCSK

Important features:

Signal peptide:

amino acids 1-29

FIGURE 261

GAGGATTTGCCACAGCAGCAGTAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTT GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTTGCAGG**ATG**ATGGTGGCCCTT CGAGGAGCTTCTGCATTGCTGGTTCTGTTCCTTGCAGCTTTTCTGCCCCCGCCGCAGTGTAC CCAGGACCCAGGCGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGG AAAAATGTACCCAAGCAACGAGGGCATACATTCAAGAATTCCAAGAGTTCTCAAAAAATATA TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTT ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC TTTGAAAATAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC ATACGGGCATTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC CTGGCAGGGAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTCATAACCAAGCAACTT CTAATGAGATAATCAAATATAACCTGCAGAAGAGGGACTGTGGAAGATCGAATGCTGCTCCCA GGAGGGTAGGCCGAGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT GGATGAGCATGGGCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAG GATGCTGAAGCCTCATTCCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC CCAACTTGTTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCCAGAGAT AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAAACTCCAGACAAAGAGAAA GCTGCCTCTGAAG<u>TAA</u>TGCATTACAGCTGTGAGAAAGAGCACTGTGGCCTTTGGCAGCTGTTC AGTGTGTAGAAGTGGAAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAA CCTCCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTTTTACT CCCTAATATTCACCACTGGCTTTTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCAACTAAAATACTATTAATATTTCTTT CTTTTCTTTTTTTTTTTGAGACAAGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACTCC AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC CACACCTGGCTTAAAATACTATTTCTTATTGAGGTTTAACCTCTATTTCCCCTAGCCCTGTC CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT CCAGGTGTGGAGTGTTTGCACATCATTGAATTCTCGTTTCACCTTTGTGAAACATGCACAAG TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT TGTTCACCTACTCTTATAGTCAATGCGTTCATCGTTTCAGCCTAAAAATAATAGTCTGTCCC TTTAGCCAGTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCCTGT

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFRVLEQGLEKCTQATRAYIQEFQE FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL QEAEEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNSPKVYLLIGSRNNTV WEFANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDT PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 263

GGGCGCCCGCGTACTCACTAGCTGAGGTGGCAGTGGTTCCACCAACATCGCAGA TGTCGGAGCTCATGGGGCTGTCGGTGTTGCTTGGGCTGCTGGCCCTGATGGCGACGGCGGCG GTAGCGCGGGGGGGCCGCCGCCGGCCGGCCGCCTGCCAAAAAGC AAATGGATTTCCACCTGACAAATCTTCGGGATCCAAGAAGCAGAAACAATATCAGCGGATTC GGAAGGAGAGCCTCAACAACACAACTTCACCCACCGCCTCCTGGCTGCAGCTCTGAAGAGC CACAGCGGGAACATATCTTGCATGGACTTTAGCAGCAATGGCAAATACCTGGCTACCTGTGC AGATGATCGCACCATCCGCATCTGGAGCACCAAGGACTTCCTGCAGCGAGAGCACCGCAGCA TGAGAGCCAACGTGGAGCTGGACCACGCCACCCTGGTGCGCTTCAGCCCTGACTGCAGAGCC TTCATCGTCTGGCTGGCCAACGGGGACACCCTCCGTGTCTTCAAGATGACCAAGCGGGAGGA TGGGGGCTACACCTTCACAGCCACCCCAGAGGACTTCCCTAAAAAGCACAAGGCGCCTGTCA TCGACATTGGCATTGCTAACACAGGGAAGTTTATCATGACTGCCTCCAGTGACACCACTGTC ACACGCTGCTGTATCTCCCTGTGGCAGATTTGTAGCCTCGTGTGGCTTCACCCCAGATGTGA AGGTTTGGGAAGTCTGCTTTGGAAAGAAGGGGGGGAGTTCCAGGAGGTGGTGCGAGCCTTCGAA CTAAAGGGCCACTCCGCGGCTGTGCACTCGTTTGCTTTCTCCAACGACTCACGGAGGATGGC TTCTGTCTCCAAGGATGGTACATGGAAACTGTGGGACACAGATGTGGAATACAAGAAGACC AGGACCCCTACTTGCTGAAGACAGGCCGCTTTGAAGAGGCGGCGGGTGCCGCCGTGCCGC CTGGCCCTCTCCCCCAACGCCCAGGTCTTGGCCTTGGCCAGTGGCAGTAGTATTCATCTCTA CAATACCCGGCGGGGGAGAAGGAGGAGTGCTTTGAGCGGGTCCATGGCGAGTGTATCGCCA ACTTGTCCTTTGACATCACTGGCCGCTTTCTGGCCTCCTGTGGGGACCGGGCGGTGCGGCTG TTTCACAACACTCCTGGCCACCGAGCCATGGTGGAGGAGATGCAGGGCCACCTGAAGCGGGC CTCCAACGAGAGCACCCGCCAGAGGCTGCAGCAGCTGACCCAGGCCCAAGAGACCCTGA AGAGCCTGGGTGCCCTGAAGAAG**TGA**CTCTGGGAGGGCCCGGCGCAGAGGATTGAGGAGGAG GGATCTGGCCTCCTCATGGCACTGCTGCCATCTTTCCTCCCAGGTGGAAGCCTTTCAGAAGG AGTCTCCTGGTTTTCTTACTGGTGGCCCTGCTTCTTCCCATTGAAACTACTCTTGTCTACTT AGGTCTCTCTTCTTGCTGGCTGTGACTCCTCCCTGACTAGTGGCCAAGGTGCTTTTCTTC CTCCCAGGCCCAGTGGGTGGAATCTGTCCCCACCTGGCACTGAGGAGAATGGTAGAGAGGAG AGGAGAGAGAGAGATGTGATTTTTGGCCTTGTGGCAGCACATCCTCACACCCAAAGAAG TTTGTAAATGTTCCAGAACAACCTAGAGAACACCTGAGTACTAAGCAGCAGTTTTGCAAGGA TGGGAGACTGGGATAGCTTCCCATCACAGAACTGTGTTCCATCAAAAAGACACTAAGGGATT TCCTTCTGGGCCTCAGTTCTATTTGTAAGATGGAGAATAATCCTCTCTGTGAACTCCTTGCA AAGATGATATGAGGCTAAGAGAATATCAAGTCCCCAGGTCTGGAAGAAAAGTAGAAAAGAGT AGTACTATTGTCCAATGTCATGAAAGTGGTAAAAGTGGGAACCAGTGTGCTTTGAAACCAAA TTAGAAACACATTCCTTGGGAAGGCAAAGTTTTCTGGGACTTGATCATACATTTTATATGGT TGGGACTTCTCTCTCGGGAGATGATATCTTGTTTAAGGAGACCTCTTTTCAGTTCATCAAG

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASCG
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AOETLKSLGALKK

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG CAGTGTTTTGCCTTCACCCCAAGTGACCATGAGGGTGCCACGCGAGTCTCAATCATGCTCC TCCTAGTAACTGTGTCTGACTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT GGGGCAGGCACCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT GGGGCGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA AGCACCACACCTGTCCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGCAGGTAC $\tt CGCTGCTCCATGGACTTGAAGAACATCAATTTT{\color{red}{\textbf{TAG}}} GCGCTTGCCTGGTCTCAGGATACCCA$ GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG CTGTGGTGTAAAGGTGGCCAGCCTGGTTCTCTCCCTGCTCAGGCTGCCAGAGAGGTGGTA AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG $\verb|CCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGT|\\$ GCATTGCTCAGAGTCCCAGGTCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC TCAAGATTGGCTCTTCCCAGAGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACTCCTTTCCA AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG TGGGAGCAAGGGACAGGGAGCAGGGCTGAAAGGGGCACTGATTCAGACCAGGGAGG CAACTACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAA

FIGURE 266

 ${\tt MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP} \\ {\tt GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF}$

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA**ATG**TCTTTC $\tt CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC$ TGCTGCCTTGGCATTGCTGCTAACACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA GCAAAGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT CCTCTGTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAGTTCTATGGTCCACAAAGGCGGAAGAT GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACTTCTTCCGAGCCTGGAACGGAG GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGGAGTTTTCGTGGTGGGATCA GGAAAGCAGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAA**TGA**T TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTCATGGGATGTATT GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA CAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC TCTACTAAAAATACAAAAATCACCCGGGTGTGGTGGCAGCCACCTGTAGTCCCAGCTACCCG GGAGGCTGAGGCAGGAGATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA

FIGURE 268

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKAALEYLEDIDLKTLEKEPR TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEGFILGGVFV VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 269

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSTSSSSDGNLRPEAITAIIV VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

FIGURE 271

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK GIVKGRNLDSRGLILGAEAWGRGVKKNT

FIGURE 273

 ${\tt GCCAGGAATAACTAGAGAGGAACA}$ GTGCCTGCTGCACCAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT TGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCT TCTACTTGGAAAAAACAAAATGAATATGGACCACCAGGCAAACTGTTTGTCCATGAGTGGG CTCACCTCCGGTGGGGAGTGTTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACTGTATG GAAAAGATTGTCAATTCTTTCCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG CAAAGTATTGATTCTGTTGAATTTTGTAACGAAAAAACCCATAATCAAGAAGCTCCAAG CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATT TTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT AAATCGAATGAATCAAGCAGCAAAACATTTCCTGCTGCAGACTGTTGAAAATGGATCCTGGG TGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAAATAAAAAGC AGTGATGAAAGAAAAAAACATCATGGCAGGATTACCCACAAATCCTCTTGGGAGGAACTTCCAT CTGCTCTGGAATTAAATATGCATTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT CCGAAGTACTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCCAGAAGTCCCTT CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA TTTCTCTCTGGGATCCCAGTGGAACAATAATGGAAAATTTCACAGTGGATGCAACTTCCAAA ATGGCCTATCTCAGTATTCCAGGAACTGCAAAGGTGGGCACTTGGGCATACAATCTTCAAGC CAAAGCGAACCCAGAAACATTAACTATTACAGTAACTTCTCGAGCAGCAAATTCTTCTGTGC CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT TGAATCACAGAATGGACATACAGAAGTTTTGGAACTTTTTGGATAATGGTGCAGGCGCTGATT CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT AGCTTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCACT GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTCAGCCGAACAGCATCCGGA GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCC ACATATTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT GCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC TACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAGTACCACCATT**TGA**ACCTTA ACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTAAAAAACAAAACAATGTAAGT TTAAGATGTCGGAAAAGGATACTTTGATTAAATAAAAACACTCATGGATATGTAAAAACTGT CAAGATTAAAATTTAATAGTTTCATTTATTTGTTATTTTGTTAAGAAATAGTGATGAAC AAAGATCCTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT

FIGURE 274

MGLFRGFVFLLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTY
LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAAKHFLLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTDGEDNTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDGVYSRYFTAYTENGRYSLKVRAHG
GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRTASGGAFVVSQV
PSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP
QANPDDIDPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 275

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG CAAGCAGCTCCGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGG GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCCTGAATGATGATGGTTCGCC GGGGGCTGCTTGCGTGGATTCCCGGGTGGTGGTTTTGCTGGTGCTCCTCTGTTGCTATCTCTGT CCTGTACATGTTGGCCTGCACCCCAAAAGGTGGCGAGGAGCAGCTGGCACTGCCCAGGGCCAACAGC CCTGTACATGTTGGCCTGCACCCCAAAAGGTGACGAGGAGCAGCTGGCACTGCCCAGGGCCAACAGC CCCACGGGAAGGAGGGGTACCAGGCCGTCCTTCAGGAGTGGAGGAGCAGCACCGCAACTACGTGA GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGGAGGAGGAGGAGCAGCTCAGGAA TGGGCAGTACCAAGCCAGCGATGCTGGCCTGGGTCTGGACAGGAGAGCCCCCCAGAGAAAACCCAG GCCGACCTCCTGGCCTTCCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG CCACAGAGTATGCAGCAGTGCCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGG CCTTACCCGCCACCCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAAGCCAT GAATCAGCCTTGGAGACCCTGAACAATCCTGCAGAAAAGCGCCCAATCACCGTCCTTACACGCCT CTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAA AGGGGACCACAAACCGAATTCAAACGGCTCATCTTATTTCGACCATTCAGCCCCCATCATGAAAGTG CTTTGAGTCCTTGCTTCTGTTTGCTCACAGTAAACTCATTGTTTAAAAGCTTCAAGACATTCAAGC TGTTGGTGTTAAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAAATTAAAACAC AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

FIGURE 276

MMMVRRGLLAWISRVVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIES
ALETLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPI
MKVKNEKLNMANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR
LNTQPGKKVFYPVLFSQYNPGIIYGHHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

FIGURE 277

 ${\tt GAAAGA} {\bf TG} {\tt TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC}$ AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC CTGGAATTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG GATTATTATATTTGGTGTGATATTTTGCATCATAGTTGCAATTGCACTACTGATTTTAT CAGGGATCTGGCAACGTAGAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA CCAAGAGCAGATCATATTTTGTTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCT TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC TCAAAATATTCTAAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG TAGTTATTGATTTAAGCATTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATT TTATCCTGTTATCACACCAACAGTTGATTATATTTTTCTGAATATCAGCCCCTAATAGGAC AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG

FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVÁFSMRKVP NREATEISHVLLCNVTQRVSFWFVVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLE FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC ENMITIENGIPSDPLDMKGGILMMPS

FIGURE 279

AACTCAAACTCCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG $\tt GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC \underline{\textbf{ATG}} \texttt{T}$ ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATACAGCTCACAGCTCTTTGG CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA GAACCCTAGTATTTCTTGAAGTTAATGGAAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCCTTAA GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC AATTTGTCTGTTACATTTCCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA ATGTGTTTACTCTCTTCCCTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG TTTCTGATTAACAGTAAATCCTAAATTCAAACTGTTAAATGACATTTTTATTTTTATGTCTC TTTGTCG

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT WNFRPLDGGPEQFVFYYHIDPFQPMSGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFQHYRKKRWAER AHKVVEIKSKEEERLNQEKKVSVYLEDTD

FIGURE 281

FIGURE 282

 ${\tt MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTTAT}\\ {\tt TAASTTARKDIPVLPKWVGDLPNGRVCP}$

FIGURE 283

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

FIGURE 285

FIGURE 286

MPVPALCILWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ IQERLHTAALPA

FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACTGTGTCCTGTTCTCCCAATCAGCTCAG GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC TAAA**TAG**GTCTTTCTCCAATGTGTCCTCCAAGCAAGATTCATCATAACTTATAGGTTCATGA TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT GTCCATAGCATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCCTACATCAGAGACTCTAGGT GCTATATAATCCAAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT TACCCTTTTTTTGGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCCTTTTTTACATTT TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT TGGAAGCTGAAAACTGAATTTAAAGAATGCTATCTTGGAAAATTGCATACGTCTGTGCAATT TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTAGAAGCAAACAATTTT AAATATATTTTGTTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT CCCACTTTGCAAACTTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTCATTGCTCA

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGDLKTQIEKLWT EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQGK WSDEACRSSKRYICEFTIPK

FIGURE 289

FIGURE 290

MKLAALLGLCVALSCSSAAAFLVGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPLKLLLS SLGIPVNHLIEGSQKCVAELGPQAVGAVKALKALLGALTVFG

FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT CCTTGGCCTCCGCAGCCGATCACATGAAGGTGGTGCCAAGTCTCCTGCTCTCCGTCCTCCTG GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCCAGCCCC TCAGAACCAGCCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGCAGCAGCTT GCCAAGGAGACTTCAAACTTCGGATTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG CAACATGGTCTTCTCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCA CAGGGCCGACTGAAACCCAGATCAAGAGGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG CCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACT GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCGTGCCTATGAATTTTCGCAATGCCTCA CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTCCCAAACT GTTTGATGAGATTAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGA AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA TTTTCGTTGTCATGTCCTCAAACTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCCTCA TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCTTTG CTGACCTTAGTGAACTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA ACAGTGATTGAAGTTGAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTAC TGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTCATTTCATGATCTATGAAG AAACCTCTGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTA**TAA**TTCAGG TACCAGCAATGGATGGCAGGGGAGAGTGTTCCTTTTGTTCTTAACTAGTTTAGGGTGTTCTC AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGGATACATTCAAAGACCCCCAGCAGATGC AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACTGATTATAGAGAAGGCTA CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA CATCCTGGGTGGGACAGACCAGGACGATGCAAGATTCCATCCCACTACTCAGAATGGCATGC TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTCATTTAATGTTTTTTGGACCATGGT TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAA

FIGURE 292

MKVVPSLLLSVLLAQVWLVPGLAPSPQSPETPAPQNQTSRVVQAPREEEEDEQEASEEKAGE
EEKAWLMASRQQLAKETSNFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQI
KRGLHLQALKPTKPGLLPSLFKGLRETLSRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYF
DTECVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLTPFDP
VFTEVDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHL
ALEDYLTTDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFSPFADLSELSA
TGRNLQVSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLF
LGRVVNPTLL

FIGURE 293

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE RPRLWVMPNHQVLLGPEEDQDHIYHPQ

FIGURE 295

TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATCAACCCAACTCAGCTTCCTGC TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCT GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAAAGCAGACTACCC AGAGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTTGGC GACGCCCAGAAAACAGCATCTTATTACTCACCCTATGGCCAGCGGGAATTCACTGCGGGATT TGTTCAGTTCAGGGTATTTAATAACGAGAGAGCCAACGCCTTGTGTGCTGGAATGAGGG TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT CCCCAGCAGTGTGGAGATTTTTCTGGTTTTGATTGGAGTGGATATGGAACTCATGTTGGTTA $\texttt{CAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT} \underline{\textbf{TGA}} \texttt{GAGTTTTGTG}$ GGAGGGAACCCAGACCTCTCCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCA GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAA

FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI YQKYPVKYGEGKCWTDNGPVIPVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAAN ALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL FYR

FIGURE 297

GCGGAGCCGCCGCCGCCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC CACGAGGCTGCCGCATCCTGCCCTCGGAACA**TTG**GGACTCGGCGCGCGAGGTGCTTGGGCCG CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCCCCATGAAAGCGCAGCC ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACACCTCCAGTGCTAACTCAAC AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACTATGCAT TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA ATACAGATTGATGCTGCCCTATCAATTAATTTTTGGTTTATTAATAGTTTAAAAACAATATTCT CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT GGGGTGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA TGCCATCTGGGCATACAAATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGT AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA CAGAAATTATACAATCAAACTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPSDH TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQN TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIG CKMYYSRRGIRYRTIDEHDAII

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG ${\tt TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCCAGCG} {\tt ATG} {\tt GCGACCCTGTGGGGAGGC}$ ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA ATGCAAATATGAAGAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG CATGTTGTCCTCAGC**TAA**TTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTCACCAACT ATAATAGAGACATTTTTAAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG TGACTTTTACTAATAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA AGCACTCTCTTTTTCACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTTG AACAACTTTTTTCAAGTCACTTTACTAAACAAACTTTTGTAAATAGACCTTACCTTCTATTT TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG ACTTTTGCACTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT CTAAAATGCCTGGTGGCTTTTCACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG CAATGCATCCTAGAACAAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG GTGTGTGTGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACACATTTG TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTTGAAACATCAA

FIGURE 300

MATLWGGLLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENSGHIYNKNIS QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIYLSILGLLLLYMVYLTL VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLARSRSRANVLNKVEYAQQRWKLQVQEQ RKSVFDRHVVLS

FIGURE 301

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT PGARFQRSHLAEAFAKAKGSGGGAGGGGSGRGLMGQIIPIYGFGIFLYILYILFKVSRIILI ILHQ

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT ${\tt GGATTTGAAAGTTGAGAGCAGCATC}{\tt TTTTGCCCACTGAAACTCATCCTGCCAGTGTTAC}$ TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG GGTGATTCAGCTCTGATGGGATGTTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT AGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGGACATCTTATGC AATGATGGCTCTCCTCCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA GGAGATTGTATTTCGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG CTTCAAGGAGTGAGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCT GGTGTTCAAGAAAACCATTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC TGTGCCACAATCCTGCTGCTCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGG GAATGCCAAAAACACAGCAAGCCTTT**TGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG TGGAGACTCTCTCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC CCAGCTGTCCTCTCTCTTTTTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCGTT GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA AAAACCAACCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHFQNRVNLVGDIFRNDGSIMLQGVRESDGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVTPAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNSLEKKSGGGMPKTQQAF

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG AAACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG GCCCTGCATCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA TGAGAATGAAAATCGCACAGGAACTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAAC TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAG TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGA TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG AGAACAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT **TAA**TGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT ATGAATGCATCAGTAGCTGAAAAAAAAAAAAAA

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYL QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLTWE ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFL EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

FIGURE 307

CCCACGCGTCCGCGCAGTCGCGAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG CGCCGGAGGACCTCGGACGCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA GCCCGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG CGTCGTGGCCATGCGGGGGCTATCGCCAGCTCGTCATCCGTCAGAAGAGGCAAGCCCGCG AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACTCTTCGGCTCCAAGAAGAGGCGCAG AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG CAGCAGTCAGGCCGAGGGTGTATCTGGGTCTGAACAAAGAAGAAGAAGAAGATCATGAAAGGCAA CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC **GTAG**CCAGTGAGGGCAAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFGSKKRRRRRP EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA MNSEGYLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGG ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG CAACCTGGATATTCTGAGACATATTTTGGGGGGGATTTCAGTGAAAAAAGTGGGGGGATCCCCT CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA GAGAGGGGGAAAGGGGACGTTTTCAATAGGAGGCAAAACTCGAGGGTGGGATCCACTGAGG AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG $\texttt{CGCGCTCCGGGCGCCTGCCGGTTTGGGGGTGTCTCCTCCCGGGGCGCT} \underline{\textbf{ATG}} \texttt{GCGGCGCTGGC}$ CGCAGCGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG CTGTCCAAGGTGCGACTGTGCGGGGGGGGGCCGGCCGGGCCGGACCGCGGCCCGGAGCCTCA GCTCAAAGGCATCGTCACCAAACTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG ACGGAAGCATCCAGGGCACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC TGAGGGACTGCTCTACAGTTCGCCGCATTTCACAGCTGAGTGTCGCTTTAAGGAGTGTGTCT TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTACCGCCAGCGTCGTTCTGGCCGGGCC TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGGAAACCGAGTTAAGAAGACCAA GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC $\texttt{ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCC} \underline{\textbf{TGA}} \texttt{AATGTAGTCCCTGGACTG}$ GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACCACCTGT

FIGURE 310

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR GPEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

FIGURE 311

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503</pre>

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRIFGLKKRRLR RQDPQLKGIVTRLYCRQGYYLQMHPDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSKTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 313

ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG $\tt CTCCCCACCCCCAAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTT$ AGAACTGGTGTGGTGTTTTCCTTTCTTTTTGAATTTCCCACAAGAGGAGGAAATTAATAATACATCTGC CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT TTTTTAAATTTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCT GGATGTTGCTGTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA**ATG** TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCT GCTTGTGGTGCTGCTCGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT GCTCCTGCAGCAACCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC ACCAACACGGCTGCTGAACCTCCATGAGAACCAAATCCAGATCATCAAGTGAACAGCTTCAAGCACTTGAG GCACTTGGAAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA ACCTCAACACTCTGGAACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTT GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACT TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAG CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACT GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACAT TTACATCACAACCCTTGGAACTGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC GAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT GAGCTGAAATGTCGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATA CAGGCATGTACACATGTATGGTGAGTAATTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA GCAACCACTACTCCTTTCCTTTCTCTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG GACCACAGATAACAATGTGGGTCCCACTCCAGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACAC CACAGAGCACAAGGTCGACAGAGAAAACCTTCACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT GGTCATTTTCTACAAGATGAGGAAGCACCATCGGCAAAAACCATCACGCCCCAACAAGGACTGTTGAAATTA TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAG CAGTTCAGTGCATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATC**TAA**AACA TTTACAGAGTTACAAAAAACAACAATCAAAAAAAAAAGACAGTTTATTAAAAATGACACAAATGACTGGGCTAA TGATCTAAAGCAGACAAAAA

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVOETOI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

FIGURE 315

GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCCCAGCTCGCCCGAGGTCCGTCGGA GGCGCCCGGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC GGGATGTCCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA CACTGAGATCAAGAGGGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA GTGGTGATCACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG AGTGGCCTTTGCTTCCAATTTCCTGGCAGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCCGCTACGTGTGGAGCCAT GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT ACTGGCAGCGAATCCGAGAGAAAGAGGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT GACTACAACCACCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA CCAGTGCACAGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAGATATGAGGAAGAAGA GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT ACGGTC TGA ATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTCTTCTCGTCATTGGAGCTCAAGTCACCAGCCACAACCAGATGAGAGGTCATCTAAGTAGCA GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATT ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA CCCAACATACCATTATAGTCTCTTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419</pre>

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 317

 $\texttt{CC} \underline{\textbf{ATG}} \texttt{GCGCTCCTGCTGCTCCTCTGCTGCGGAGTAGTGGATTTCGCCAGAAGTTTGAGTATCACTACT}$ CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA ${\tt CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT}$ ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACTGTCAGATATTGGCACATATCAGTGCAAAGTGAA AAAAGCTCCTGGTGTTGCAAATAAGAAGATTCATCTGGTAGTTCTTGTTAAGCCTTCAGGTGCGAGATGTTACG TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG TATGAGTGGCAAAAATTGTCTGACTCACAGAAAATGCCCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG GCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA GTACCAAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGACTCTCCCACCTGCTAAGTTCAAGTACCCTTACAA $\texttt{GACTGATGGAATTACAGTTGTA} \underline{\textbf{TAA}} \texttt{ATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTTGACTTTATTTT}$ AGGCCTCTAGTAAAGACTTAAATGTTTTTTAAAAAAAGCACAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC ATCTACTTTATGCAATGGCATTAGACATGTAAGTCAGATGTCATGTCAAAATTAGTACGAGCCAAATTCTTTGT TAAAAAACCCTATGTATAGTGACACTGATAGTTAAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA ATTTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTTCAA AGGAAATTTTAAAATTCTTACGTTCTGTTTAATGTTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG TTCTTTTCCCCTTTTATGCACACAGAACACGCGTTGTCATGCCTCAAACTATTTTTTATTTGCAACTACA TAAAGTAAATTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGC ACAATATAAATACTCAAGTCCAATATTAAAAACTTAGGCACTTGACTAACTTTAATAAAATTTCTCAAACTA ${\tt TATCAATATCTAAAGTGCATATTTTTTTAAGAAAGATTATTCTCAATAACTTCTATAAAAATAAGTTTGATGG}$ TTTGGCCCATCTAACTTCACTACTATTAGTAAGAACTTTTAACTTTTAATGTGTAGTAAGGTTTATTCTACCTT TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTTTCTTCCC TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCTGGCTT TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGTGGAAATATAAGTTCTGAAATCTGTAGGGAAGAACACAT TAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT GAGATCACATCTAAGTGACACTCCTATTGTCAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCCTAAACAA $\verb|CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGGTTT|\\$ CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT AGCAAACTAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT TATGAACACAAAGAAGGAAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAAGGAAAAGGAGCA GAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247, 245-251, 296-302

Myelin P0 protein.

amino acids 96-125

FIGURE 319

CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGC CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGACACGAGACTG AGAG**ATG**AATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC $\tt CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG$ CTTCTCTGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT GAAGGGGGTTGTTCCCCAGAAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACTGTTTTCAAAAA CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAACT TTGTTCTCATCGTGTCACAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC AGTGCACACAGGCGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC TC<u>TGAATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCCTGTGTCATTTCA</u> AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATTC TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG GTTAAAAAACAGAGGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTCATT TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA СТССАААААААААААААААААААА

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMQMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLVHTLLEFYLKTVFKNH HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL TKALGEVDILLTWMQKFYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

FIGURE 321

AAGGAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACCACGGTCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAAATCTTGAGAAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACTCTGCGGCAATGTCAGGAACAGAGCCAGCTGCCAGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGAAGACCCCCTTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAAGAATCATCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAAGCCAAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTTGAATAAAATTCCATATTTTACCTATGA

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282</pre>

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ RQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT AGAATAACATCCACTTTGCCTTTCTCCACAGGTGTCCACTCCCAGGTCCAACTGCACCTC GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCGCCT CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCGTCCTCAGAGCCTATCCCA ATGCCTCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA TGAGCAGAAGATACCTCTGCATGGATTTCAGAGGCAACATTTTTGGATCACACTATTTCGAC CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC CACCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACC CCCATACCACGGCGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT GCTGAAGCCCCGGCCCGGATGACCCCGGCCCCGGCCTCTGTTCACAGGAGCTCCCGAGCG CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC ACGCACGCTGGGGGAACGGCCCGGAAGGCTGCCGCCCCTTCGCCAAGTTCATC**TAG**GGTCG CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARNSYHLQIHKNGHVD GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY DVYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSE RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFA KFI

Important features of the protein:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 325

 ${\tt GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG} {\color{red} {\bf ATG}}$ GGCTGTCTCTGGGGTCTGGCCCCTTTTCTTCTTCTTCTGGGGAGGTTGGGGTCTCTGG GAGCTCTGCAGGCCCCAGCACCCGCAGAGCAGACACTGCGATGACAACGGACGACACAGAAG TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACTTC CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC ${\tt GAGGCCGACCACAGACG} {\tt TGA} {\tt GTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG}$ TGGCTTCCCCGGAAGACCTCACTGACCCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC CGGGAACTCCACGCCCCACGCCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG GGCCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTTACCCAGAAGGTTCCCATG AAGGGCAGCATGTCCAAGCCCCTAACCCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCCAGAGGTGTCCTTGGACTCACCTTGG CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA TTAAAATTCACTCAGTGTGGCCCAAAAAAA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLSA
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAIFDTLCTDDSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTSAKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site. amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169, 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339, 404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 327

GCGGAGCATCCGCTGCGGTCCTCGCCGAGACCCCCGCGCGGATTCGCCGGTCCTTCCCGCGG GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCCTCTCGACGCCA CCTTGACCTTTGAAGACCAAAACTAAACTGAAATTTAAA**ATG**TTCTTCGGGGGGAGAAGGGAG CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC AAAAAACATATCAGGGGACAAAGCATGTAACTTGATGATCTTCGACACTCGAAAAACAGCTA GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG CCAAGAGTTACCCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCCTCTGATCAAGAAA TAGCTCATCTGCTGAAAATGTGAGTGCGCTCCCAGCTACGGTGGCAGTTGCTTCTCCA CATACCACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC AGCCTCCCACGACCCTCATTTCTACAGTTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA AACCATACCGTTTACAGAAATCTCCAACTTAACTTTGAACACAGGGAATGTGTATAACCCTA CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAAACTGCTTCCTGGGAAGGT AGGGAGGCCAGTCCAGGCAGTTCCTCCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCCTGTTCCTGGTGATAGGCC TCGTCCTCCTGGGTAGAATCCTTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT TATTTGATCAATGGGATCTATGTGGACATC**TAA**GGATGGAACTCGGTGTCTCTTAATTCATT TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG GCTCTGTTGCCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA CCACACCTGGGTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG GTCTCAAACTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGAGAAGGAATGAAGTG GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT ATGTAAAGTAATAAGTATAATTGCCATATAAATTTCAAAATTCAACTGGCTTTTATGCAAA GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC TGTTTATATCAATTTCTAATGGATTTGCTTTTTTTTATATGGATTCCTTTAAAACTTATT CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGIRGNEPVYTSTQED CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP TDSKGSLETIPFTEISNLTLNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 329

CTCCCACGGTGTCCAGCGCCCAGA**ATG**CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT GGGATCCTCTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGACCCAGGAGACAAT GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA ACCTCACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG TCTTTACTGATCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCTCCCCTCCCCTTCTCCCAC CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACCAGCCAAGCAGGGGAAGACAGGG GCTGAGGCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGAGCCTTCTGTCAGC CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC CCTTCCCAGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA GCTGGGCTTCTCGAAGTTTGTCTCAGCG**TAG**GGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT GAAGCAGTATGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGCATGTTCCA GCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC TGGCGTCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG GACTCTGAATTCTAACAATGCCCAGTGACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTG ATGAACGCTCACACCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC CAATAGATCTGCTCTGCCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA AAACCTTGGCTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTGAAAGTA GCACAACTACTATTTTTTTTTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT TTTGTACTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGAC CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG TCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAAACTTGGAAAG ATGGAGGAGAAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT TATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTCACATAATTTGCCGGTGTTCTT TTTACAGAGCAATTATCTTGTATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC GCTGCATAAAAAAAAAAAAA

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196</pre>

<subunit 1 of 1, 332 aa, 1 stop</pre>

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFV FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128